TRƯỜNG ĐẠI HỌC KINH TẾ QUỐC DÂN



BÀI TẬP LỚN

PHÂN TÍCH CHUỖI THỜI GIAN TRONG TÀI CHÍNH

GIẢNG VIÊN: THS.BÙI DƯƠNG HẢI LỚP TÍN CHỈ: TOTC1124(222)_01

STT	Họ tên	Đóng góp	Ðiểm
6	Nông Quốc Chính	40%	
7	Đỗ Huy Đức	60%	
	Tổng	100%	

HÀ NỘI, 2023

Mục lục

1. Giới thiệu chung về ngành công nghệ thông tin trong thị trường chứng khoán Việ	et Nam 1
2. Phân tích chuỗi thời gian với mã chứng khoán ngành công nghệ thông tin	1
2.1. Nông Quốc Chính - Chuỗi số liệu về lợi nhuận sau thuế của CTCP Công trình Viett Giới thiệu mã chứng khoán CTR Phân tích lợi nhuận sau thuế Chuỗi số liệu về giá cố phiếu. Kết luận chung	1 2 4
2.2. Đỗ Huy Đức - Chuỗi số liệu về lợi nhuận sau thuế của CTCP Tập đoàn Công nghệ	CMC theo
quý	7 7 9
3.Phân tích đồng tích hợp 2 chuỗi CMG và CTR	12
3.1.Lý thuyết đồng tích hợp	12
3.2.Phân tích hồi quy giả mạo	12
3.3. Phân tích đồng tích hợp	13
3.4.Mô hình hiệu chỉnh sai số	14
4. Sử dụng mô hình VAR cho 2 chuỗi CMG và CTR	15
4.1.Lý thuyết mô hình VAR	15
4.2.Dự báo VAR	16
4.3.Hàm phản ứng	17
4.4.Phân rã phương sai	18
5. Kết luận	19
6. Tham khảo	21

1. Giới thiệu chung về ngành công nghệ thông tin trong thị trường chứng khoán Việt Nam

Viễn thông được hiểu là lĩnh vực chuyên sâu về nghiên cứu và sử dụng các thiết bị, tạo nên mạng lưới truyền dẫn thông tin, phục vụ cho mục đích giao tiếp xuyên biên giới. Song song viễn thông là khái niệm điện tử. Cụ thể, lĩnh vực Điện tử nghiên cứu và chế tạo ra các vi mạch điện tử, cũng chính là "bô não" của các thiết bi thông minh, điều khiển toàn bô hoat đông.

Từ đó, ngành điện tử - viễn thông ra đời. Có thể hiểu đơn giản, ngành điện tử viễn thông ứng dụng công nghệ tiên tiến, tạo ra các thiết bị truyền dẫn thông tin. Những sản phẩm quen thuộc đã được đưa vào trong đời sống hàng ngày có thể kể đến như tivi, điện thoại, máy tính, mạch điều khiển,...

Ngành viễn thông nói riêng và ngành điện tử - viễn thông nói chung giữ vai trò quan trọng trong việc xây dựng hệ thống thông tin liên lạc toàn cầu. Từ đó, quá trình trao đổi, truy xuất thông tin, giám sát và điều khiển thiết bị thông minh diễn ra nhanh chóng và thuận tiện hơn.

Việt Nam được đánh giá là một quốc gia có nhiều lợi thế để ngành viễn thông có nhiều tiến bộ hơn nữa trong tương lai. Sự tác động của cuộc cách mạng công nghệ 4.0 đã làm thay đổi mọi mặt, mọi lĩnh vực của đời sống. Không nằm ngoài xu thế đó, ngành viễn thông chịu ảnh hưởng của khoa học - công nghệ cũng có nhiều sự chuyển mình, đột phá. Trong thời đại công nghiệp hóa, hiện đại hóa như hiện nay, ngành điện tử viễn thông được dự đoán sẽ phát triển theo những phương hướng sau đây.

2. Phân tích chuỗi thời gian với mã chứng khoán ngành công nghệ thông tin

2.1. Nông Quốc Chính - Chuỗi số liệu về lợi nhuận sau thuế của CTCP Công trình Viettel theo quý

Giới thiệu mã chứng khoán CTR

Tổng Công tr CP Công trình Viettel, tiền thân là Xí nghiệp Khảo sát thiết kế và Xí nghiệp Xây lắp Công trình được thành lập ngày 30 tháng 10 năm 1995, là đơn vị hạch toán độc lập, một phần không thể thiếu trong hệ sinh thái Tập đoàn Công nghiệp - Viễn thông Quân đội. Sau khi làm tốt các nhiệm vụ của đất nước, người Công trình Viettel đã ý thức và bắt tay vào hành động mang đến một cuộc sống chú trọng trải nghiệm, sự tiện nghi, hiện đại hơn đến với công chúng.

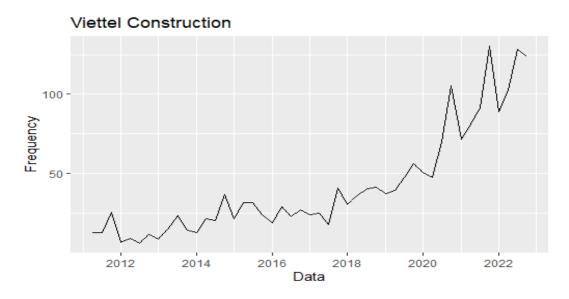
Sau nhiều lần chuyển đổi tên giao dịch, đến tháng 9/2018, Công trình Viettel chính thức mang tên Tổng Công ty CP Công trình Viettel (Tên giao dịch quốc tế Viettel Construction Joint Stock Corporation) hoạt động kinh doanh với 6 trụ chính: Đầu tư Hạ tầng, Xây dựng (bao gồm xây dựng hạ tầng viễn thông, xây dựng dân dụng B2C và B2B), Công nghệ thông tin, Giải pháp Tích hợp, Dịch vụ Kỹ thuật và Vân hành Khai thác.

Trải qua 27 năm phát triển và đồng hành cùng công chúng, Tổng Công ty đã xây dựng được hệ thống hạ tầng mạng lưới viễn thông trong nước với hơn 50.000 trạm phát sóng, 320.000 km cáp

quang (tương đương 8 vòng trái đất), bao phủ 100% lãnh thổ Việt Nam bao gồm vùng biển hải đảo, vùng sâu vùng xa. Đồng thời Tổng Công ty còn tiên phong tham gia xây dựng hạ tầng viễn thông tại 10 quốc gia ở châu Á, châu Phi và châu Mỹ. Từ đơn vị xây lắp thuần túy, trở thành Tổng Công ty cung cấp dịch vụ trên nền tảng kĩ thuật, công nghệ với quy mô gần 11.000 người, tâm đắc mang trong mình sứ mệnh đồng hành "Dựng Xây Cuộc Sống Mới" cùng các doanh nghiệp trên cả nước và thế giới.

Phân tích lợi nhuận sau thuế

Phân tích số liệu của CTCP Công trình Viettel có mã chứng khoán CTR, cụ thể ở đây là về lợi nhuận sau thuế (tính theo quý) lấy từ năm 2011 đến năm 2022, và giá đóng cửa của cổ phiếu tình từ đầu năm 2021 đến cuối năm 2022



Hình 2.1.2 Lợi nhuận sau thuế của CTR

Đây là đồ thị tổng quan về lợi nhuận sau thuế của công ty tính từ quý 1 năm 2011 đến quý 4 năm 2022. Lợi nhuận cao thường rơi vào quý 4 với quý 3 hàng năm. Tuy nhiên trong và sau đại dịch Covid, lợi nhuận tăng trưởng con số, thậm chí còn có những thời điểm tăng gấp đôi so với cùng kỳ năm trước.Điều này sẽ ảnh hưởng đến chỉ số MAPE của chuỗi.

Ta sẽ sử dụng một vài các mô hình đơn giản (hồi quy theo xu thế thời gian) để dự báo cho lợi nhuận gộp các quý trong năm 2023 của công ty.

Mô hình Trend – season – add.

```
> #Trend, Seas, Add
> model1 = lm(Data~trend + s1 + s2+ s3, data = train)
> mape(train$Data, model1$fitted.values) #full
[1] 0.4016996
> mape(test$Data, predict(model1, test)) #test
[1] 0.3018175
> predict(model1, forecast)
1 2 3 4
74.08795 81.02799 84.37035 96.11962
```

Mô hình Trend – season – multi:

Mô hình Holt-winter add

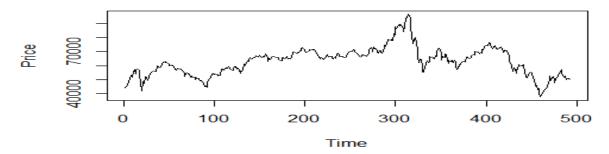
```
> #Holt-winter, add
> nr = ts(train$Data, start = c(2, 2011), frequency = 4)
> hwa = HoltWinters(nr, seasonal="a")
> mape(train[5:43,]$Data, hwa$fitted[,1])#full
[1] 0.26804
> mape(test$Data, predict(hwa, 4)) #test
[1] 0.125717
> predict(hwa, 8)
Qtrl Qtr2 Qtr3 Qtr4
515 102.3536 110.9833 125.9057
516 155.9101 131.9509 140.5807 155.5030 517 185.5074
```

Mô hình Holt-winter multi

Những mô hình trên đều được dự báo với training set từ năm 2011 đến năm 2021. Ta có thể thấy rằng MAPE (2022) và MAPE tổng thể bé nhất rơi vào mô hình Holt-winter multi. Vì thế, mô hình này sẽ được chọn để dự báo cho năm 2023. Hầu như MAPE của tất cả 4 mô hình phía trên đều nhỏ, điều này được giải thích bằng việc lợi nhuận gộp của công ty hàng quý tang trưởng ổn định so với kỳ trước. Lợi nhuận cao thường được thấy ở quý 3 và quý 4, cũng có những quý lợi nhuận trung bình.

Chuỗi số liệu về giá cố phiếu.

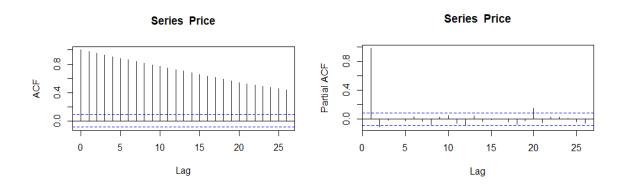
Ta sẽ sử dụng chuỗi giá cổ phiếu của công ty với mã chứng khoán CTR từ 4/1/2021 đến 29/12/2022.



Hình 2.1.3 Line plot của chuỗi giá CTR

Trên là mô hình tổng quan về giá cổ phiếu của công ty trong hai năm 2021 và 2022. Nhìn chung giá cổ phiếu tăng trưởng lành mạnh do nhiều tiêu điểm đầu từ nhưng tụt thê thảm ở một số thời điểm (cụ thể là do đứt cáp)

Ta sẽ xác định BCE có phải là chuỗi dừng hay không qua hai đồ thị ACF và PACF.

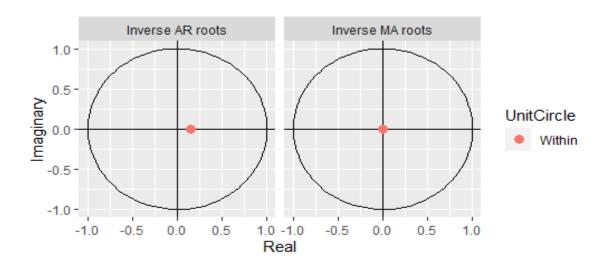


Hình 2.1.4 ACF và PACF

Theo hai đồ thị, ta kết luận chuỗi CTR là một chuỗi không dừng. Sau đó ta sử dụng ADF để phân tích thêm về tính dừng của chuỗi sai phân. Ta nhận được giá trị τ_q s =-14,49 so với τ_0 ,01 =-3,98. Bác bỏ H_0 , không có nghiệm đơn vị, ta xác định đây là chuỗi sai phân là chuỗi dừng.

Tiếp đến, ta xác định được bậc ARIMA của chuỗi là (1,1,1), bằng cách sử dụng vòng lặp để tìm ra mô hình ARIMA với AIC nhỏ nhất.

* Đánh giá tính dừng của mô hình qua vòng tròn đơn vị



Hình 2.1.5 Vòng tròn đơn vị của ARIMA(1,1,1)

Các nghiệm nghịch đảo đều nằm trong vòng tròn đơn vị => mô hình cho kết quả dừng.

* Đánh giá xem u_t còn tự tương quan không, có phải là nhiễu trắng không

Ljung-Box test

```
data: Residuals from ARIMA(1,1,1) with drift Q^* = 7.5229, df = 8, p-value = 0.4814 Model df: 2. Total lags used: 10
```

 H_0 : Không tự tương quan

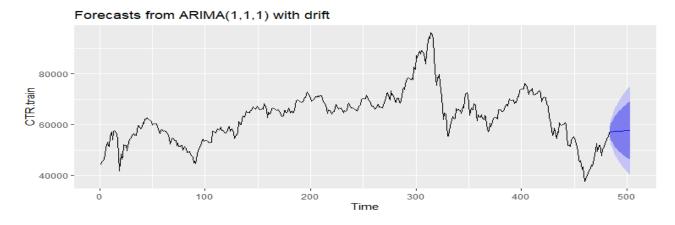
 H_1 : Có tự tương quan

P-value > 0.05. Chưa bác bỏ H_0 . u_t , là nhiễu trắng.

Cuối cùng là dự báo giá cổ phiếu của 10 ngày đầu tiên năm 2023. Ta có kết quả:

57483.97; 57511.45; 57538.94; 57566.43; 57593.91; 57621.40; 57621.40; 57648.88; 57676.37; 57703.86; 57731.34

Với RMSE và MAPE có kết quả lần lượt là 0.01994554và 1713.624, đồng thời MAPE của 10 giá trị dự báo tiếp theo là 0.1123819. Mô hình ARIMA (1,1,1) có thể sử dụng để dự báo được .



Hình 2.1.6 Dự báo mô hình ARIMA

Bên trên là đồ thị dự báo của mô hình ARIMA (1,1,1).

Kết luận chung

Từ các kết quả ước lượng cũng như dự báo từ 4/1/2021 đến 29/12/2022 ta có thể thấy được có một sự chênh lệch nhất định so với thực tế, tuy nhiên kết quả khá chính xác. Như vậy cũng có thể thấy được mô hình ARMA đã giải thích được sự biến động của mã cổ phiếu CTR. Theo kết quả của dự báo thì giá đóng cửa của CAN có xu ổn định. Công ty về Công ty Cổ phần Công trình Viettel là một công ty cổ phần nên các nhà đầu tư muốn trở thành cổ đông dài hạn của CTR có thể sử dụng ARMA làm một trong những phương pháp để đưa ra quyết định đầu tư của mình. Như vậy, ta có thể nhận định rằng mô hình ARMA khá phù hợp khi sử dụng trong ngắn hạn mặc dù trong thực tế có rất nhiều tác động dẫn đến nhiều sai số khác.

2.2. Đỗ Huy Đức - Chuỗi số liệu về lợi nhuận sau thuế của CTCP Tập đoàn Công nghệ CMC theo quý

Giới thiệu mã chứng khoán CMG

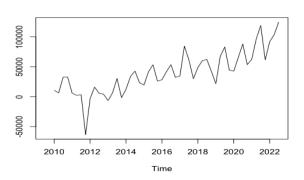
Công ty Cổ phần Tập đoàn Công nghệ CMC tiền thân là Trung tâm ADCOM thuộc Viện Công nghệ vi Điện tử, Viện nghiên cứu Công nghệ Quốc gia thành lập năm 1991 với số vốn ban đầu 50 triệu đồng. Ngày 7/2/2007, CMC thực hiện cổ phần hóa chính thức chuyển đổi thành Công ty Cổ phần Tập đoàn Công nghệ CMC. Hiện nay, vốn điều lệ của công ty đã lên tới hơn 1.089 nghìn tỉ đồng, cùng với đó là sự tham gia đóng góp của hơn 6000 cán bộ nhân viên. CMG cũng được biết tới là một công ty lớn trong ngành công nghệ thông tin tại Việt Nam khi có tới 108 triệu cổ phiếu đang được lưu hành. Hiện nay, các cổ đông chính của công ty CMC gồm Samsung SDS Asia Pacific Pte.Ltd, Công ty TNHH đầu tư MVI và Tập đoàn Geleximco với cổ phần lần lượt là 29.88%,13.5% và 10.01%.

Phân tích lợi nhuận sau thuế của công ty CMC

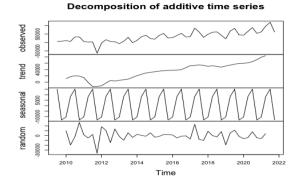
Để phân tích lợi nhuận gộp theo quý của mã CMG, ta sẽ sử dụng 8 mô hình hồi quy tuyến tính bằng việc sử dụng dữ liệu chuỗi thời gian là lợi nhuận gộp hằng quý của CMG tính từ quý 3 năm 2009 cho tới quý 4 năm 2021 để train, còn lại ta sẽ sử dụng dữ liệu trong năm 2022 để test.

Do trong dữ liệu, quý 4 năm 2009 không được ghi nhận, ta sẽ sử dụng lệnh na_kalman() để ước lượng hợp lý tối đa cho dữ liệu tại thời gian quý 4 năm 2009.

Nhìn vào đồ thị, ta có thể thấy được lợi nhuận sau thuế của công ty CMC có xu hướng tăng trong giai đoạn từ 2012 cho tới hiện tại. Tuy nhiên tại năm cuối năm 2011, chúng ta thấy được một giá trị ngoại lại, tương ứng với shock khi mà trong 2 quý liên tiếp, CMC ghi nhận lợi nhuận sau thuế âm, đỉnh điểm là quý 4 năm 2011 với khoản lỗ lên tới 63000 triệu đồng. Giải thích cho sự thua lỗ này, CMC đã chỉ ra rằng trong năm 2011 thua lỗ do tình hình kinh tế khó khăn, dự báo chiến lược hạn chế, lạm phát tăng đột biến cũng như bộ phận CMC Blue France tại Pháp hoạt động kinh doanh không hiệu quả đã gây ra khoản lỗ khổng lồ này. Sau khi đã tái cấu trúc và có những chiến lược kinh doanh phù hợp, lợi nhuận sau thuế của CMC đã tăng trở lại và có xu hướng tăng nhanh dần.



Hình 2.2.1: Lợi nhuận sau thuế của CMG



Hình 2.2.2: Phân tách thành phần dạng cộng chuỗi lợi nhuận sau thuế CMG

Qua đồ thị, ta sẽ thấy rõ được chuỗi có tính mùa vụ khi hai quý đầu luôn thấp hơn 2 quý sau, từ đó ta có thể thấy được doanh thu chủ yếu của CMC sẽ luôn đến vào 2 nửa cuối của năm.

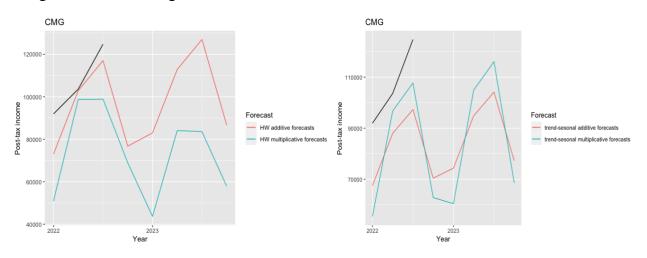
Sử dụng mô hình hồi quy để ước lượng cho các hệ số của mô hình tương ứng, từ đó ta sẽ lựa chọn mô hình dựa trên R^2 và MAPE

Mô hình	\mathbb{R}^2	MAPE ₂₀₂₂	MAPE	Dự báo năm 2023 (t = 55:59)
$RE_t = \beta_0 + B_1 t + \epsilon_t$	0.61	0.242	0.792	$RE_{55} = 84946.35$ $Re_{56} = 86650.77$
				$RE_{57} = 88355.19$ $RE_{58} = 90059.62$
$RE_t = \beta_0 + \beta_1 \ln(t) + \epsilon_t$	0.41	0.456	1.691	$RE_{55} = 58418.38$ $Re_{56} = 58830.78$
				$RE_{57} = 59235.89 RE_{58} = 59633.94$
$RE_t = \beta_0 + \beta_1 t + \alpha_1 s 1 + \alpha_2 s 2 + \alpha_3 s 3$	0.75	0.212	0.859	$RE_{55} = 74370.05$ $Re_{56} = 94860.72$
$+ \epsilon_t$				$RE_{57} = 104154.7$ $RE_{58} = 77282.74$
$RE_t = \beta_0 + B_1 t * s1 + \beta_2 * t * s2 + B_3 * t$	0.78	0.199	0.792	$RE_{55} = 60436.67$ $Re_{56} = 104963.3$
$*s3 + \epsilon_t$				$RE_{57} = 116049.2$ $RE_{58} = 68598.82$
$RE_{t+h t} = l_t + hb_t + s_{t+h-m(k+1)}$		0.091	1.139	$RE_{55} = 82985.45$ $Re_{56} = 112871.89$
				$RE_{57} = 126982.5$ $RE_{58} = 86702.46$
$RE_{t+h t} = (l_t + hb_t)s_{t+h-m(k+1)}$		0.23	1.128	$RE_{55} = 43671.62$ $Re_{56} = 84088.16$
, , ,				$RE_{57} = 83617.65$ $RE_{58} = 57863.97$

Bảng 2.2.3: Dự báo lợi nhuận sau thuế bằng các mô hình hồi quy

Có thể thấy được sự chênh lệch giữa MAPE của toàn bộ dữ liệu và MAPE₂₀₂₂ rất lớn, nó cho thấy được rằng mô hình trên không phù hợp cho việc phân tích dữ liệu theo quý của công ty. Có một điểm đáng chú ý đó là mô hình dự báo Holt-Winter dạng cộng lại cho thấy được MAPE₂₀₂₂ chỉ có 9% trong khi MAPE đầy đủ của cả mô hình là 113%. Lí giải cho nguyên nhân này, đó là do trong những năm cuối cùng, số liệu đã được làm tron nên xảy ra sự chênh lệch tương đối lớn so với MAPE đầy đủ khi trong những năm 2011 công ty.

Chính vì vậy, việc vẽ đồ thị dự báo lợi nhuận sau thuế của CMG bằng những mô hình trên chỉ mang tính chất ước lượng.



Hình 2.2.4 : Dự báo Holt-Winter

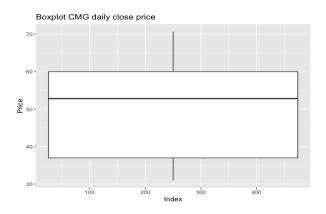
Hình 2.2.5: Dự báo Trend-Seasonal

Giới thiệu cổ phiếu CMG và phân tích chung

Trong giai đoạn 2021-2022, mặc dù dưới sự ảnh hưởng của Covid-19 khiến cả thị trường chứng khoán giảm điểm, dường như nhóm ngành công nghệ thông tin nói chung và mã CMG nói riêng dừng như có sự bứt phá ngoạn mục trong giai đoạn này khi giá cổ phiếu đã tăng tới 40% giá trị khi đã đạt mốc 40.000 đồng/cổ phiếu. Nguyên nhân dẫn tới đà tăng này là do sự phát triển của công nghệ mạng 5G cũng như sự thúc đẩy về ứng dụng công nghệ thông tin của chính phủ tại Việt Nam đã tạo điều kiện thuận lợi để các công ty công nghệ bứt phá. Trong năm 2021, CMC đã đẩy mạnh dịch vụ lưu trữ Cloud cũng như mở rộng Data Center là thế mạnh lớn nhất của công ty.

Trong 4 tháng đầu năm 2022, dưới sự tác động của các gói kích thích kinh tế của chính phủ khi thuế VAT đã được giảm còn 8% cùng với sự phục hồi của thị trường, VnIndex đã thủng ở ngưỡng 1500 điểm với sự gồng gánh chủ yếu của ngành BDS và ngân hàng đã kéo theo sự phát triển của các nhóm ngành khác như CNTT. Trong đà tăng trưởng của cả thị trường, CMG cũng đã tạo đỉnh mới khi đạt mức giá kỉ lục 70.700 đồng và mức giá của CMG được giữ ở ngưỡng 60000 cho tới tháng 8 cùng năm. Sự kiện Agribank thoái toàn bộ vốn lên tới 3% cổ phần khỏi CMC đã tạo ra đà giảm cho mã này khi nó đã kéo CMC về lại mức giá 32.000 đồng và chỉ phục hồi trở lại khi công ty thông báo trả cổ tức.

Để phân tích cho chuỗi CMG,ta sử dụng 499 quan sát từ dữ liệu giá đóng cửa của mã chứng khoán CMC bắt đầu từ 1/1/2021 tới 31/12/2022. Số lượng dữ liệu trên sẽ được sử dụng thống kê mô tả và kiểm định bằng việc sử dụng mô hình ARIMA và so sánh dự báo cho 12 phiên đầu năm 2023 và giá đóng cửa thực tế của 12 phiên đầu năm 2023.



Min. 1:	st Qu.	Median
Mean 3:	rd Qu.	Max.
31.00	36.98	52.80
49.50	60.00	70.70

Với kết quả thống kê mô tả như trên, ta sẽ thấy được giá thấp nhất và cao nhất của CMG có sự chênh lệch tới hơn 100% và chuỗi giá luôn ổn định ở mức trung bình với giá đạt 49.500 dồng.

Hình 2.2.6: Boxplot của giá đóng cửa mã CMG

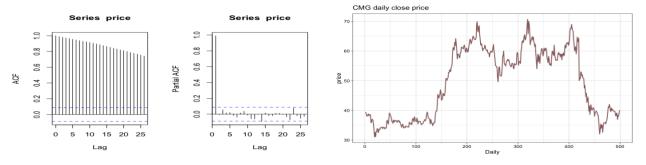
Phân tích giá đóng cửa của CMG bằng sử dụng mô hình ARIMA

Mô hình ARIMA(p,d,q) có thể được biểu diễn như sau.

$$\phi(L)(1-L)^dY_t = \mu + \theta(L)u_t$$

Để ước lượng mô hình ARIMA cho chuỗi, ta sẽ áp dụng phương pháp Box-Jenkins với chuỗi giá đóng cửa của CMG.

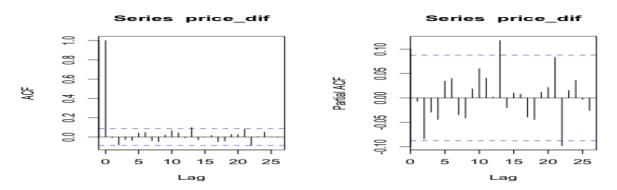
Nhìn vào đồ thị ACF và PACF của chuỗi CMG, ta dễ dàng thấy được chuỗi là bước ngẫu nhiên khi mà ACF cho thấy được sự tương quan của các thời kì liên tục với nhau cùng với kiểm định Augmented Dicky-Fuller, cho ra kết quả thống kê $|\tau|=0.37<|\tau_{0.01}|=2.58$, kết quả kiểm định trên càng làm rõ rằng chuỗi giá CMG không dừng. Để khiến chuỗi dừng, ta sẽ lấy sai phân bậc 1 của chuỗi và sẽ có kết quả như hình.



Hình 2.2.7: ACF và PACF của chuỗi giá CMG

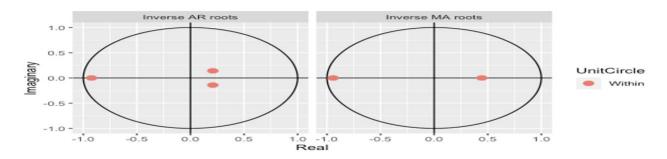
Hình 2.2.8: Line plot của chuỗi giá CMG

Qua đồ thị ACF và PACF của chuỗi sai phân bậc 1, ta sẽ dễ dàng thấy được chuỗi đã dừng và từ đó lựa chọn MA(2) từ ACF và AR(3) từ PACF do chúng đã gần vượt qua mức tin cậy 95%. Sau khi lựa chọn ARMA(3,2) ta lựa chọn Intergrated bằng kiểm định Dicky-Fuller với hệ số chặn và xu thế và thu được kết quả $|\tau_3|$, $|\phi_2|$, $|\phi_3|$ đều lớn hơn mức ý nghĩa 1%, cho thấy hệ số chặn, xu thế có ý nghĩa. Một mô hình khác cũng có thể sử dụng được là ARIMA(13,1,13) khi các lag đã thoà mãn.



Hình 2.2.8: ACF và PACF của chuỗi Δ_{CMG}

Nhìn vào vòng tròn đơn vị, ta thấy được các nghiệm AR và MA đều nằm trong vòng tròn đơn vị, từ đó ta đánh giá được mô hình ARIMA(3,1,2) phù hợp với việc dự báo cho chuỗi CMG.



Hình 2.2.9: Vòng tròn đơn vị của ARIMA(3,1,2)

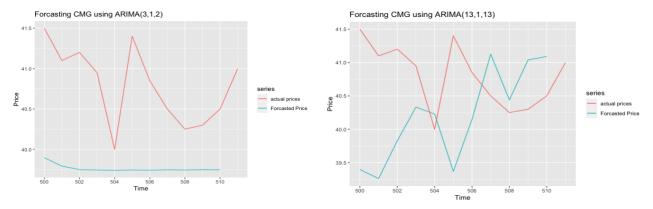
Từ các kiểm định trên, chúng ta có thể viết mô hình ARIMA(3,1,2) như sau:

$$\Delta CMG_t - \mu^* = \phi_1(\Delta CMG_{t-1} - \mu^*) + \phi_2(\Delta CMG_{t-2} - \mu^*) + \phi_3(\Delta CMG_{t-3} - \mu^*) + \mu_t + \theta_1\mu_{t-1} + \theta_2\mu_{t-2}$$

Mô hình trên đưa ra AIC = 1824.02, BIC = 1853.5 chúng ta sẽ sử dụng tiêu chỉ này để so sánh với ARIMA(13,1,13) ở dưới.

Với mô hình ARIMA(3,1,2) với hệ số chặn, ta sử dụng kiểm định Ljung-Box và có được kết quả Với kết quả kiểm định trên, ta bác bỏ H0 và chấp nhận H1 là phần dư được phân phối độc lập và μ_t là nhiễu trắng với p-value = 0.2899

Dự báo mô hình ARIMA(3,1,2), ta thu được kết quả sau



Hình: 2.2.10: Dự báo mô hình ARIMA(3,1,2) Hình 2.2.11: Dự báo mô hình ARIMA(13,1,13)

Với sự chênh lệch trên,mô hình ARIMA(3,1,2) không nên được để dự báo cho giá của CMG.Tương tự với mô hình ARIMA(13,1,13), ta có kết quả dự báo sau:

Lặp lại các bước trên, với kiểm định Ljung-Box, p-value ARIMA(13,1,13) = 0,012 cũng như tất cả các nghiệm của mô hình nằm trong vòng tròn đơn vị phù hợp với mô hình. Nhưng một lần nữa khi dự báo cho 10 phiên đầu lại không cho kết quả chính xác. AIC lần này đưa ra với kết quả là 1830.11 và BIC =1948.01 cao hơn một chút so với ARIMA(3,1,2). Dựa trên tiêu chí AIC,BIC cũng như các lý thuyết kinh tế, mô hình ARIMA(3,1,2) sẽ được ưu tiên lựa chọn hơn, tuy nhiên như đã

phân tích ở trên, con số dự báo này không chính xác nên chúng ta không sử dụng mô hình để dự báo được.

3. Phân tích đồng tích hợp 2 chuỗi CMG và CTR

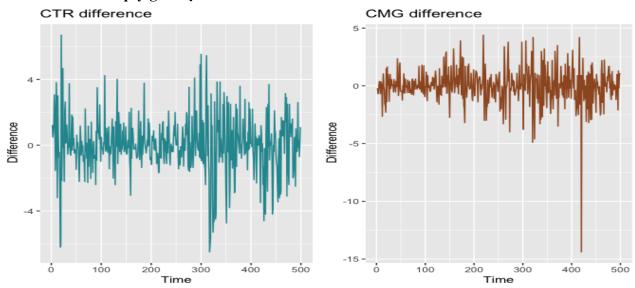
3.1.Lý thuyết đồng tích hợp

Hồi quy giả mạo (Spurious regression) là khi hai chuỗi không có quan hệ, cùng không dừng, cùng xu thế, hồi quy có ý nghĩa thống kê. Trong trường hợp 2 chuỗi CMG và CTR, 2 chuỗi trên có quan hệ với nhau là đều là 2 mã cổ phiếu đầu ngành nhóm CNTT tại thị trường chứng khoán Việt Nam, vậy khả năng 2 chuỗi là hồi quy giả mạo thấp.

Hai chuỗi thời gian được gọi là đồng tích hợp có nghĩa là tồn tại một số nguyên nhân gốc làm cho hai chuỗi thời gian có xu hướng biến động tương đồng nhau về dài hạn ngay cả khi nếu nhìn riêng lẻ thì chúng là random walks.

Chúng ta kiểm định đồng tích hợp với 2 chuỗi CMG và CTR bởi 2 chuỗi này đều thuộc nhóm ngành CNTT, do đó chúng có khả năng đồng tích hợp với nhau.

3.2.Phân tích hồi quy giả mạo



Hình 3.1: Sai phân bậc 1 của CTR và CMG

Như đã thực hiện ở bài cá nhân, 2 chuỗi sai phân của CTR và CMG đều không dừng. Với CTR, độ biến động sai phân bé hơn rất nhiều số với CMG khi tại t =413, việc Agribank thoái vốn khỏi CMG đã làm giá cổ phiếu này mất hơn 10% giá trị của nó đã gây ra shock cho mã này.

Tiếp theo, chúng ta sẽ ước lượng phương trình xem cân bằng dài hạn có xảy ra không.

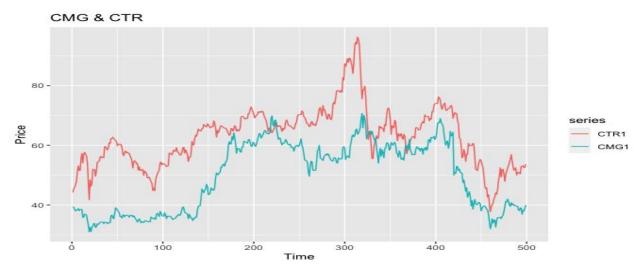
```
Call:
                                                   Call:
lm(formula = CMG1 ~ CTR1)
                                                   lm(formula = CTR1 ~ CMG1)
Residuals:
                                                   Residuals:
    Min
            10 Median
                               30
                                                      Min
                                                               10 Median
                                                                                30
                                                                                          Max
                                       Max
-17.3302 -6.3685 0.3441
                           5.7038 21.7564
                                                   -17.3347 -3.9232 -0.5112 3.0674 24.0883
Coefficients:
         Estimate Std. Error t value Pr(>|t|)
                                                   Coefficients:
(Intercept) -6.79322 2.12985 -3.19 0.0151 **
                                                            Estimate Std. Error t value Pr(>|t|)
CTR1
           0.89569
                      0.03348 26.76 <2e-16 ***
                                                   (Intercept) 30.22733 1.25194 24.14 <2e-16 ***
                                                                       0.02463 26.76
                                                          0.65899
                                                                                          <2e-16 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05
                                                   Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
'.' 0.1 ' ' 1
                                                   0.1 ' ' 1
Residual standard error: 7.427 on 497 degrees of
                                                   Residual standard error: 6.371 on 497 degrees of
Multiple R-squared: 0.5902, Adjusted R-squared:
                                                  Multiple R-squared: 0.5902, Adjusted R-squared:
F-statistic: 715.9 on 1 and 497 DF,
p-value: < 2.2e-16
                                                  F-statistic: 715.9 on 1 and 497 DF,
                                                  p-value: < 2.2e-16
```

Với kết quả kiểm định như trên, CMG và CTR đều phụ thuộc lẫn nhau ở ngưỡng p-value thoả mãn ở mức 5%, điều này cho thấy rằng lý thuyết mô hình đã được thoả mãn. Để xác định cân bằng dài hạn, ta cần kiểm định xem phần dư có phải là nhiễu trắng hay không.

Phần dư của 2 mô hình hồi quy trên đều là nhiễu trắng, vì vậy chúng có cân bằng dài hạn cho nhau, vậy nên khả năng xảy ra đồng tích hợp rất cao.

3.3. Phân tích đồng tích hợp

Nhận xét đồ thị gốc



Hình 3.2: Chuỗi giá đóng cừa của CMG và CTR

Qua đồ thị của 2 biến CMG và CTR, cả 2 chuỗi đều có xu hướng chậm trong giữa năm 2021, thời điểm Covid-19 ở Việt Nam bắt đầu hoành hành. Tại thời điểm t=300 tức đầu năm 2022 khi VNIndex liên tục lập kỉ lục, khi đó 2 chuỗi có xu hướng tăng rõ rệt nhất, với đà đi xuống của VNIndex trong những năm tiếp theo, CMG và CTR không tránh khỏi được sự suy thoái của thị trường khi liên tục mất giá và chỉ thực sự hồi phụ vào cuối năm 2022. Chính bởi đà tăng giảm có vẻ đồng đều, ta đưa ra giả thuyết sau :

Kiểm định đồng tích hợp

Kiểm định đồng tích hợp (cointergrated test) là kiểm định giúp kiểm tra mối quan hệ tương ứng của biến X với biến Y trong chuỗi thời gian. Kiểm định này sẽ thể hiện X có thể được biểu diễn theo Y không. Với 2 chuỗi CMG và CTR, ta đưa ra 2 cặp giả thuyết sau:

Cặp giả thuyết 1. H0: Số mối quan hệ đồng tích hợp = 0 H1: Số mối quan hệ đồng tích hợp > 0 Cặp giả thuyết 2.

H0: Số mối quan hệ đồng tích hợp ≤ 1 H1: Số mối quan hệ đồng tích hợp > 1

Sử dụng kiểm định Johansen với tiêu chí "trace" để kiểm tra có đồng tích hợp giữa 2 mã CMG và CTR không, ta có kết quả kiểm định như sau Cặp giả thuyết 1: Bác bỏ H0 ở độ tin cậy 5%, chấp nhận H1, có đồng tích hợp xảy ra trong chuỗi Cặp giả thuyết 2: Chưa có cơ sở bác bỏ H0, chấp nhận H0, chỉ có dưới 1 cặp đồng tích hợp trong chuỗi

Với kết quả trên, có thể thấy rằng có đồng tích hợp xảy ra giữa 2 chuỗi CMG và CTR. Điều này thể hiện rằng 2 chuỗi trên có xu thế giống nhau ở mức tin cậy 5%.

3.4.Mô hình hiệu chỉnh sai số

Mô hình ECM thể hiện thay đổi trong CMG phụ thuộc vào sự thay đổi trong CTR và độ trễ của hạng nhiễu cân bằng u_{t-1}. Nếu hạng nhiễu này bằng không, thì sẽ không có sự mất cân bằng nào giữa hai biến và trong trường hợp đó mối quan hệ dài hạn sẽ được xác định bởi mối liên hệ đồng liên kết (không có hạng nhiễu ở đây). Nhưng nếu hạng nhiễu cân bằng khác không, thì mối quan hệ giữa CMG và CTR sẽ chệch ra ngoài trạng thái cân bằng. Phương trình hiệu chỉnh sai số sẽ cho chúng ta biết được tác động tức thời của biến CTR lên biến CMG.

$$\Delta CMG_t = (\alpha_0) + \alpha_1 \Delta CTR_t + \gamma u_{t-1} + v_t$$

```
Call:
                                                            Có thể thấy phần dư có ý
lm(formula = dif CMG ~ dif CTR +
                                                            nghĩa thống kê ở mức ý
residCMG[1:498])
                                                            nghĩa 1%, và hệ số của
Coefficients:
                                                            phần dư bé hơn 0, vậy nên
                Estimate Std. Error t value
                                                            quan hệ giữa CTR và
Pr(>|t|)
(Intercept) -6.201e-05 6.622e-02 -0.001
                                                            CMG sẽ lệch ra khỏi trạng
0.999253
                                                            thái cân bằng. Đồng thời
dif CTR
               7.091e-02 3.848e-02 1.843
                                                            tác đông tức thời của CTR
0.065945 .
                                                            lên CMG là mất cân bằng
residCMG[1:498] -3.263e-02 8.931e-03 -3.653
0.000287 ***
                                                            của kì trước.
Signif. codes: 0 '***' 0.001 '**' 0.01 '*'
0.05 \.' 0.1 \ ' 1
Residual standard error: 1.478 on 495 degrees
of freedom
Multiple R-squared: 0.03183, Adjusted R-
squared: 0.02792
F-statistic: 8.138 on 2 and 495 DF, p-value:
0.0003331
```

4. Sử dụng mô hình VAR cho 2 chuỗi CMG và CTR

4.1.Lý thuyết mô hình VAR

Mô hình VAR hay còn gọi là mô hình vectơ tự hồi quy là một dạng tổng quát của mô hình tự hồi quy đơn chiều (*univariate autoregressive model*) trong dự báo một tập hợp biến, nghĩa là một vector của biến chuỗi thời gian. Nó ước lượng từng phương trình của mỗi biến chuỗi theo các độ trễ của biến (p) và tất cả các biến còn lại (Vế phải của mỗi phương trình bao gồm một hằng số và các độ trễ của tất cả các biến trong hệ thống). Một cách đơn giản, mô hình VAR 2 chiều với 1 độ trễ có dạng hệ 2 phương trình như sau

$$Y_t = \beta_{10} + \beta_{11}Y_{t-1} + \alpha_{11}X_{t-1} + u_{1t}$$
$$X_t = \beta_{20} + \beta_{21}Y_{t-1} + \alpha_{21}X_{t-1} + u_{2t}$$

Trong mô hình VAR, mối quan hệ 2 chiều của biến được thể hiện rõ ràng với nhau. Mô hình VAR tổng quát hoá mô hình tự động hồi quy đơn biến bằng cách cho phép sử dụng chuỗi thời gian đa biến. Mô hình VAR kế thừa tính chất của mô hình AR khi các giá trị của biến bao gồm giá trị trễ đều bị ảnh hưởng bởi sự biến động của shock trong quá khứ. Trong VAR, không còn biến phụ thuộc và biến độc lập riêng biệt, mô hình không cần lý thuyết kinh tế chi tiết mà chỉ cần ảnh hưởng lẫn trong trong chuỗi thời gian.

VAR sẽ được sử dụng cho chuỗi dừng để dự báo ảnh hưởng qua lại lẫn nhau của biến. Trong trường hợp chuỗi gốc không dừng mà xảy ra hiện tượng đồng tích hợp, điều đó cho thấy chuỗi kết hợp tuyến tính của các chuỗi trên là chuỗi dừng. Trong trường hợp này ,VAR sẽ thêm vào cơ chế làm tron sai số và sử dụng ước lượng bình phương nhỏ nhất để ước lượng.

4.2.Dự báo VAR

Với 2 chuỗi CTR và CMG, ta sẽ kiểm định hệ nhân quả Granger trước để tìm xem 2 biến có giải thích cho nhau không, từ đó ta sẽ áp dụng mô hình VAR(1), từ đó dự báo cho 2 chuỗi trên.

$$CMG_{t} = \beta_{10} + \left(\beta_{11}CMG_{t-1} + \dots + \beta_{1p}CMG_{t-p}\right) + \left(\alpha_{11}CTR_{t-1} + \dots + \alpha_{1p}CTR_{t-p}\right) + u_{1t}$$

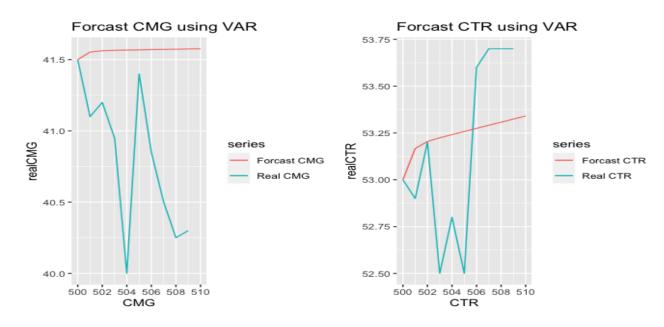
$$CTR_{t} = \beta_{20} + \left(\beta_{21}CTR_{t-1} + \dots + \beta_{2p}CTR_{t-p}\right) + \left(\alpha_{21}CMG_{t-1} + \dots + \alpha_{2p}CMG_{t-p}\right) + u_{2t}$$

Kiểm định cho Δ_{CMG} giải thích cho Δ_{CTR}	Kiểm định cho Δ_{CTR} giải thích cho Δ_{CMG}		
H0: $\alpha_{11} = \alpha_{12} = \alpha_{1p}$	H0: $\beta_{11} = \beta_{12} \dots = \beta_{1p}$		
H1: Hệ số $\alpha \neq 0$	H1: Hệ số $\beta \neq 0$		
P-value = 0.8915	P-value = 0.1268		

Với kết quả kiểm định trên, ta thấy rằng Δ_{CMG} không hề giải thích cho Δ_{CTR} và ngược lại. Bởi dữ liệu chỉ có 2 chuỗi nên ta sẽ sử dụng mô hình VAR(1). Áp dụng mô hình VAR(1), ta có kết quả sau

```
VAR Estimation Results:
                                                         Trong mô hình ước lượng thứ
_____
                                                         nhất, chuỗi CMG đã không
Estimation results for equation dif CMG:
_____
                                                         phu thuôc theo CTR, shock
dif CMG = dif CMG.11 + dif CTR.11 + const
                                                         của CTR, trong khi đó với mô
          Estimate Std. Error t value Pr(>|t|)
                                                         hình ước lượng thứ hai, chuỗi
dif CMG.11 -0.0121643 0.0450347 -0.270 0.787
                                                         CTR có ý nghĩa với shock và
dif CTR.11 0.0598864 0.0391529 1.530
                                      0.127
        0.0005856 0.0672756 0.009
const
                                      0.993
                                                         chính nó trong quá khứ
                                                         nhưng không có phản ứng với
Residual standard error: 1.5 on 494 degrees of freedom
                                                         chuỗi CMG.
Multiple R-Squared: 0.004762, Adjusted R-squared: 0.0007331
F-statistic: 1.182 on 2 and 494 DF, p-value: 0.3075
Estimation results for equation dif CTR:
______
dif CTR = dif CMG.11 + dif CTR.11 + const
         Estimate Std. Error t value Pr(>|t|)
dif CMG.11 -0.007005 0.051341 -0.136 0.89153
dif CTR.11 0.145587 0.044636 3.262 0.00118 **
const 0.014169 0.076696 0.185 0.85351
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1
Residual standard error: 1.71 on 494 degrees of freedom
Multiple R-Squared: 0.0211, Adjusted R-squared: 0.01714
F-statistic: 5.325 on 2 and 494 DF, p-value: 0.005151
```

Trong kiểm định tự tương quan của phần dư của mô hình VAR(1), ta có p-value = 0.0867 > 0.05, ta bác bỏ H0 và chấp nhận H1 là phần dư không có tự tương quan. Với việc phần dư không có tự tương quan, mô hình VAR(1) có thể sử dụng để dự báo được.



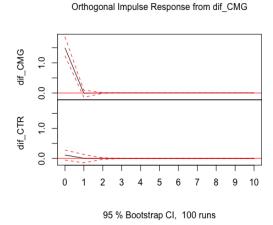
Hình 4.1: Dự báo giá đóng cửa CMG và CTR theo VAR(1)

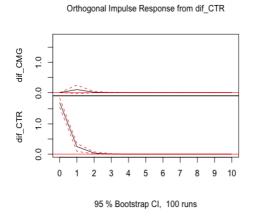
Với kết quả dự báo như trên, ta có thể thấy rằng sai phân của 2 chuỗi CMG và CTR được dự báo sẽ ổn định trong thời gian dài, trong khoảng tin cậy 95% của 2 chuỗi đều ngưỡng xấp xỉ 3, con số này sẽ không hợp lý trong dự báo khi phương sai là quá lớn trong ngắn hạn.

4.3. Hàm phản ứng

Hàm phản ứng sử dụng các mô hình VAR để đánh giá sự biến động của các biến trong mô hình với các shock bởi một hoặc nhiều biến khác nhau. Chính bởi vậy, hàm phản ứng sẽ đánh giá được tác động đơn lẻ của các shock tới mô hình. Với CMG và CTR, ta có đồ thị hàm phản ứng như hình.

Nhìn vào hàm phản ứng của biến sai phân CMG, ta thấy rõ được tác động của shock chính nó lên sự biến động của CMG trong khi CTR gần như có ảnh hưởng quá nhiều trong ngắn hạn với giá trị xấp xỉ 0. Với hàm phản ứng của CTR, điều tương tự cũng xảy ra khi tác động của shock CMG lên CTR gần như bằng 0 trong khi CTR chịu ảnh hưởng chỉnh bởi shock của chính nó.





Hình 4.2: Hàm phản ứng của Δ_{CMG}

Hình 4.3: Hàm phản ứng của Δ_{CTR}

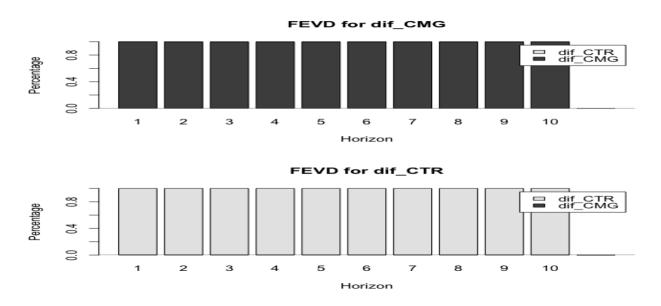
4.4.Phân rã phương sai

Phân rã phương sai là một cách tiếp cận để phân tích cấu trúc mô hình VAR. Phân rã phương sai phân tích sự biến thiên của một biến do tác động bởi cú sốc của chính biến đó và cú sốc của các biến nội sinh khác. Chúng ta sẽ tác các thành phần của biến để xem chính nó và biến khác giải thích được bao nhiêu phần trăm cho chính nó ở các thời kì trước. Phân rã phương sai được tính theo công thức sau:

$$\sum_{i} v_{jih} = 100\%, \ \forall j, h$$

Với v_{jih} là tỉ lệ (%) MSE của biến Y_j được giải thích bởi sốc (biến động) của Y_i , tính cho h kì. Tuy nhiên với việc hàm phản ứng của Δ_{CMG} và Δ_{CTR} thể hiện ở trên chỉ có tác động của Δ_{CMG} và Δ_{CTR} với shock chính nó thời kì trước, ta có thể suy đoán rằng % biến động của shock của Δ_{CMG} và Δ_{CTR} chỉ ảnh hưởng lên chính chuỗi gốc, không có tác động lên mã còn lại.

Sau khi sử dụng mô hình VAR cho 2 biến CMG và CTR, ta sẽ thực hiện phân tích phương sai và ra kết quả sau:



Hình 4.4: Hàm phân rã phương sai của Δ_{CMG} và Δ_{CTR}

Đồ thị phân rã phương sai của 2 biến sai phân CMG và CTR cho thấy được hầu như chỉ có sự giải thích của chính nó trong các thời kì trước lên hiện tại, tác động của biến còn lại là không đáng kể. Việc tự giải thích cho chính mình mặc dù 2 biến trên cùng nhóm ngành CNTT càng làm rõ hơn sự độc lập của 2 công ty khi mà CTR tập trung vào quản trị, xây dựng viễn thông và CMG với mảng kinh doanh chính là lưu trữ Cloud.

5. Kết luận

Sau khi phân tích 2 cổ phiếu CMG và CTR thuộc nhóm ngành CNTT của thị trường chứng khoán Việt Nam dựa trên 2 chuỗi lợi nhuận sau thuế và giá đóng cửa cổ phiếu, ta có những nhận xét sau

Thứ nhất, lợi nhuận của nhóm ngành CNTT tăng trưởng đều đặn và có xu hướng tăng dần trong những năm trở lại đây với sự bùng nổ của cách mạng công nghiệp 4.0 tại thế giới cũng như quy trình số hoá tại Việt Nam. Các mô hình dự báo tăng trưởng như Holt-Winter, Trend-Seasonal của lợi nhuận sau thuế đều không hợp lý khi dự báo có sự chênh lệch lớn với thực tế.

Thứ hai, về dự báo giá với mô hình ARIMA, cả hai mô hình đều cho kết quả dự báo lệch so với giá thực tế, do vậy không thể sử dụng ARIMA để dự báo cho giá cổ phiếu của 2 mô hình trên. Giải thích cho điều này có thể do mô hình ARIMA chưa bao gồm đủ yếu tố tác động với sự thay đổi về giá cổ phiếu của 2 công ty trên, hơn nữa sai số còn phụ thuộc vào nhiều yếu tố bên ngoài tính kinh tế.

Thứ ba, hai chuôi CMG và CTR có đồng tích hợp với nhau ở độ tin cậy 5%. Điều này cho thấy rằng chúng ta có thể dự báo được hướng tăng, giảm của CMG theo CTR hoặc ngược lại. Tác động

tức thời của 2 chuỗi qua lại với nhau chỉ diễn ra trong ngắn hạn và sẽ bị chệch khỏi cân bằng trong dài hạn khi 2 công ty có chiến lược phát triển và sản phẩm chính khác nhau.

Thứ tư, sau khi áp dụng mô hình VAR(1), ta chỉ thấy được sự tác động của chính bản thân chuỗi và shock của chính nó trong các thời kì trước chứ không phụ thuộc vào mã còn lại. Kết quả dự báo của mô hình VAR với 2 chuỗi trên đều chỉ mang tính chất thử nghiệm chứ không thể áp dụng vào dự báo thực tế khi chúng có quá nhiều sai số.

Thứ năm, như đã dự đoán ở kiểm định nhân quả Granger, khi phân rã phương sai và kiểm tra hàm phản ứng từ mô hình VAR, CMG cũng chỉ bị ảnh hưởng bởi chính nó và shock của chính mình trong các thời kì trước, ngược lại với CTR

Cuối cùng, qua bài dự báo trên, ta có thể thấy được rằng hầu hết kết quả dự báo cho lợi nhuận sau thuế và giá đóng cửa của cổ phiếu sẽ có xu hướng tăng nhưng kết quả dự báo không chính xác, không thể sử dụng các mô hình trên để dự báo vào thực tiễn. Bởi lẽ thị trường chứng khoán Việt Nam là một thị trường phức tạp với rất nhiều nhân tố ảnh hưởng không chỉ bởi chính nó và shock ở những thời kì trước mà tại đây, ảnh hưởng của các nhân tố ngoài hay sai số là rất lớn, không thể lường trước được. Trong nhóm ngành CNTT nói riêng, do tiềm năng phát triển còn rất lớn cũng như đang được nhà nước mọi điều kiện để thúc đẩy phát triển do nên chúng ta cần thần trọng khi đầu tư vào ngành này.

6. Tham khảo

- 1. https://otexts.com/fpp2/
- 2. https://www.vietlod.com/mo-hinh-var
- 3. https://www.r-econometrics.com/timeseries/irf/
- 4. https://www.uef.edu.vn/newsimg/tap-chi-uef/2014-03-04-15/10.pdf
- 5. https://finance.vietstock.vn/CMG-ctcp-tap-doan-cong-nghe-cmc.htm
- 6. https://finance.vietstock.vn/CTR-tong-cong-ty-co-phan-cong-trinh-viettel.htm
- 7. https://s.cafef.vn/tin-doanh-nghiep/cmg/Event.chn