

ĐẠI HỌC KINH TẾ QUỐC DÂN
KHOA TOÁN KINH TẾ



BÀI TẬP NHÓM
PHÂN TÍCH CHUỖI THỜI GIAN VÀ DỰ BÁO

**ĐỀ BÀI: PHÂN TÍCH VÀ DỰ BÁO CHUỖI GIÁ CỔ PHIẾU
VÀ CHUỖI DOANH THU**

Lớp học phần: TOKT1147(224)_01

GV hướng dẫn: Bùi Dương Hải

Nhóm thực hiện: Nhóm 05

| Sinh viên: | MSV |
|--------------------|----------|
| Phạm Hạnh Nguyên | 11224830 |
| Trần Thị Kiều Oanh | 11225075 |
| Nghiêm Gia Phương | 11225223 |
| Nguyễn Thị Trà | 11226303 |

Hà Nội, 2025

Mục lục

| | |
|--|------------|
| DANH MỤC BẢNG BIỂU | v |
| DANH MỤC HÌNH | vii |
| 1 LỜI MỞ ĐẦU | 1 |
| 2 TỔNG QUAN | 2 |
| 2.1 Tổng quan các nhóm ngành nghiên cứu | 2 |
| 2.1.1 Tổng quan ngành ngân hàng | 2 |
| 2.1.2 Tổng quan ngành dầu khí | 2 |
| 2.2 Tổng quan về dữ liệu nghiên cứu | 3 |
| 2.2.1 Thu thập số liệu | 3 |
| 2.2.2 Xử lý dữ liệu | 3 |
| 3 BÀI CÁ NHÂN | 4 |
| 3.1 Phân tích chuỗi VCB - Phạm Hạnh Nguyên | 5 |
| 3.1.1 Tổng quan về Ngân hàng TMCP Ngoại thương Việt Nam | 5 |
| 3.1.2 Dự báo chuỗi doanh thu năm 2025 | 5 |
| 3.1.3 Dự báo giá cổ phiếu 10 phiên đầu năm 2025 | 7 |
| 3.2 Phân tích chuỗi CTG - Trần Thị Kiều Oanh | 11 |
| 3.2.1 Tổng quan về ngân hàng Viettinbank | 11 |
| 3.2.2 Dự báo chuỗi doanh thu năm 2025 | 11 |
| 3.2.3 Dự báo giá cổ phiếu 10 phiên đầu năm 2025 | 13 |
| 3.3 Phân tích chuỗi PVC - Nghiêm Gia Phương | 17 |
| 3.3.1 Tổng quan về Tổng Công ty Hóa chất và Dịch vụ dầu khí - CTCP | 17 |
| 3.3.2 Dự báo doanh thu cho 4 quý năm 2025 của PVC | 17 |
| 3.3.3 Dự báo giá cổ phiếu PVC 10 phiên đầu năm 2025 | 19 |
| 3.3.4 Dự báo rủi ro cho 10 phiên đầu năm 2025 bằng mô hình GARCH | 22 |
| 3.4 Phân tích chuỗi PVS - Nguyễn Thị Trà | 23 |
| 3.4.1 Giới thiệu | 23 |
| 3.4.2 Dự báo chuỗi doanh thu năm 2025 | 23 |
| 3.4.3 Dự báo giá cổ phiếu 10 phiên đầu năm 2025 | 25 |
| 3.4.4 Mô hình ARCH/GARCH và các dự báo | 28 |
| 4 BÀI NHÓM | 29 |
| 4.1 Xây dựng danh mục tối ưu | 29 |
| 4.1.1 Phân tích tổng quan | 29 |
| 4.1.2 Tối thiểu hóa phương sai | 30 |
| 4.1.3 Tối đa hóa lợi nhuận | 31 |
| 4.2 Dự báo chuỗi lợi suất danh mục | 33 |

| | | |
|-----------------|---|-----------|
| 4.2.1 | Từ chuỗi lợi suất gốc của danh mục | 33 |
| 4.2.2 | Từ kết quả chuỗi dự báo cá nhân | 36 |
| 4.2.3 | So sánh kết quả dự báo lợi suất chuỗi portfolio từ 2 cách | 36 |
| 4.3 | Phân tích đồng tích hợp | 37 |
| 4.4 | Phân tích VAR | 38 |
| 4.4.1 | Xác định bậc của trễ | 38 |
| 4.4.2 | Ước lượng mô hình VAR(1) | 39 |
| 4.4.3 | Kiểm định tương quan phần dư | 40 |
| 4.4.4 | Dự báo bằng mô hình VAR(1) | 40 |
| 4.4.5 | Phân tích hàm phản ứng và phân rã phương sai | 43 |
| 5 | KẾT LUẬN | 48 |
| ĐÁNH GIÁ | | 49 |
| PHỤ LỤC | | 50 |

Danh sách bảng

| | | |
|------|--|----|
| 3.1 | Bảng kết quả dự báo cho chuỗi doanh thu VCB trong năm 2024 | 6 |
| 3.2 | Bảng kết quả dự báo cho chuỗi doanh thu VCB trong năm 2025 | 7 |
| 3.3 | Bảng kết quả kiểm định tính dừng chuỗi VCB ở mức ý nghĩa 5% | 8 |
| 3.4 | Bảng so sánh các mô hình ARIMA cho chuỗi sai phân ΔVCB_t | 9 |
| 3.5 | Bảng so sánh các mô hình ARIMA cho chuỗi log - return VCB | 10 |
| 3.6 | Dự báo giá cổ phiếu VCB 10 phiên đầu năm 2025 | 10 |
| 3.7 | Bảng kết quả dự báo cho chuỗi doanh thu CTG trong năm 2024 | 12 |
| 3.8 | Bảng kết quả dự báo cho chuỗi doanh thu CTG trong năm 2025 | 13 |
| 3.9 | Bảng kết quả kiểm định tính dừng chuỗi CTG ở mức ý nghĩa 5% | 14 |
| 3.10 | So sánh các mô hình ARIMA cho chuỗi giá đóng cửa CTG | 15 |
| 3.11 | So sánh mô hình ARIMA(0,0,4) và ARIMA(2,0,2) | 16 |
| 3.12 | So sánh hiệu suất dự báo với Giá thực | 16 |
| 3.13 | Các mô hình dự báo doanh thu của PVC | 18 |
| 3.14 | Kết quả dự báo doanh thu 4 quý năm 2025 của PVC | 19 |
| 3.15 | Bảng kết quả kiểm định tính dừng ở mức ý nghĩa 5% | 20 |
| 3.16 | So sánh các mô hình ARIMA | 21 |
| 3.17 | Dự báo giá cổ phiếu PVC 10 phiên đầu năm 2025 | 22 |
| 3.18 | Dự báo rủi ro của giá cổ phiếu PVC 10 phiên đầu năm 2025 | 22 |
| 3.19 | Bảng kết quả dự báo cho chuỗi doanh thu PVS trong năm 2024 | 24 |
| 3.20 | Kết quả dự báo cho chuỗi doanh thu PVS trong năm 2025 | 24 |
| 3.21 | Kết quả kiểm định tính dừng của chuỗi giá PVS | 25 |
| 3.22 | Kết quả kiểm định tính dừng của chuỗi sai phân giá PVS | 25 |
| 3.23 | Kết quả kiểm định tính dừng của chuỗi log-return giá PVS | 26 |
| 3.24 | Kết quả mô hình ARIMA cho chuỗi giá đóng cửa PVS | 26 |
| 3.25 | Kết quả mô hình ARIMA cho chuỗi log-return giá PVS | 27 |
| 3.26 | So sánh kết quả chuỗi giá được tính từ mô hình cho chuỗi giá và log-return | 28 |
| 3.27 | Dự báo rủi ro cho 10 phiên đầu năm 2025 | 28 |
| 4.1 | Bảng giá trị trung bình của giá cổ phiếu và lợi suất các chuỗi | 29 |
| 4.2 | Kết quả phương án tối thiểu hóa phương sai danh mục | 31 |
| 4.3 | Kết quả phương án tối đa hóa lợi suất kì vọng danh mục | 32 |
| 4.4 | Bảng kết quả kiểm định tính dừng của chuỗi lợi suất danh mục | 34 |
| 4.5 | Bảng so sánh các mô hình ARIMA cho chuỗi lợi suất danh mục | 35 |
| 4.6 | Log-return dự báo, giá quy đổi và giá thực tế của danh mục trong 10 phiên đầu năm 2025 | 36 |
| 4.7 | Chuỗi log-return dự báo của các cổ phiếu và portfolio theo phiên | 36 |
| 4.8 | Dữ liệu chuỗi lợi suất thực | 37 |
| 4.9 | So sánh hai cách theo phiên | 37 |
| 4.10 | Kiểm định Johansen giữa các chuỗi giá | 38 |
| 4.11 | Hồi quy VAR cho r_{VCB} | 39 |
| 4.12 | Hồi quy VAR cho r_{CTG} | 39 |

| | | |
|------|--|-----|
| 4.13 | Hồi quy VAR cho r_{PVC} | 39 |
| 4.14 | Hồi quy VAR cho r_{PVS} | 40 |
| 4.15 | Dự báo giá cổ phiếu VCB 10 phiên đầu năm 2025 | 41 |
| 4.16 | Dự báo giá cổ phiếu CTG 10 phiên đầu năm 2025 | 42 |
| 4.17 | Dự báo giá cổ phiếu PVC 10 phiên đầu năm 2025 | 42 |
| 4.18 | Dự báo giá cổ phiếu PVS 10 phiên đầu năm 2025 | 43 |
| 5.1 | Kết quả hồi quy lin - lin | 50 |
| 5.2 | Kết quả hồi quy lin - log | 51 |
| 5.3 | Kết quả hồi quy log - lin | 51 |
| 5.4 | Kết quả hồi quy log - log | 52 |
| 5.5 | Kết quả hồi quy với biến giả mùa vụ | 52 |
| 5.6 | Kết quả xu thế tuyến tính có mùa vụ, dạng cộng | 53 |
| 5.7 | Kết quả xu thế tuyến tính có mùa vụ, dạng nhân | 53 |
| 5.8 | Kết quả Holt - Winter có mùa vụ, dạng cộng | 54 |
| 5.9 | Kết quả Holt - Winter có mùa vụ, dạng nhân | 54 |
| 5.10 | Holt - Winter có mùa vụ, dạng cộng; sử dụng để dự báo cho 2025 | 55 |
| 5.11 | Kiểm định ADF cho chuỗi giá của VCB, có xu thế | 56 |
| 5.12 | Kiểm định ADF cho chuỗi sai phân của VCB, có xu thế | 57 |
| 5.13 | Kiểm định ADF cho chuỗi lợi suất của VCB, có xu thế | 58 |
| 5.14 | Kiểm định ADF cho chuỗi giá của VCB, có hệ số chặn | 59 |
| 5.15 | Kiểm định ADF cho chuỗi sai phân của VCB, có hệ số chặn | 60 |
| 5.16 | Kiểm định ADF cho chuỗi lợi suất của VCB, có hệ số chặn | 61 |
| 5.17 | Kiểm định ADF cho chuỗi giá của VCB, không có hệ số chặn | 62 |
| 5.18 | Kiểm định ADF cho chuỗi sai phân của VCB, không có hệ số chặn | 63 |
| 5.19 | Kiểm định ADF cho chuỗi lợi suất của VCB, không có hệ số chặn | 64 |
| 5.20 | Mô hình ARIMA(3,1,3) cho chuỗi sai phân VCB | 64 |
| 5.21 | Mô hình ARIMA(1,1,3) cho chuỗi sai phân VCB | 65 |
| 5.22 | Mô hình ARIMA(3,1,1) cho chuỗi sai phân VCB | 65 |
| 5.23 | Mô hình ARIMA(3,1,0) cho chuỗi sai phân VCB | 66 |
| 5.24 | Mô hình ARIMA(3,0,3) cho chuỗi lợi suất VCB | 67 |
| 5.25 | Mô hình ARIMA(1,0,3) cho chuỗi lợi suất VCB | 68 |
| 5.26 | Mô hình ARIMA(3,0,1) cho chuỗi lợi suất VCB | 68 |
| 5.27 | Mô hình ARIMA(3,0,0) cho chuỗi lợi suất VCB | 69 |
| 5.28 | Kết quả hồi quy lin - lin | 71 |
| 5.29 | Kết quả hồi quy lin - log | 71 |
| 5.30 | Kết quả hồi quy log - lin | 72 |
| 5.31 | Kết quả hồi quy log - log | 72 |
| 5.32 | Kết quả hồi quy với biến giả mùa vụ | 73 |
| 5.33 | Kết quả xu thế tuyến tính có mùa vụ, dạng cộng | 73 |
| 5.34 | Kết quả xu thế tuyến tính có mùa vụ dạng nhân | 74 |
| 5.35 | Kết quả xu thế không tuyến tính có mùa vụ dạng cộng | 74 |
| 5.36 | Kết quả xu thế không tuyến tính có mùa vụ dạng nhân | 75 |
| 5.37 | Kết quả Holt-Winter có mùa vụ dạng cộng | 75 |
| 5.38 | Kết quả Holt-Winter có mùa vụ dạng nhân | 76 |
| 5.39 | Kiểm định ADF cho chuỗi lợi suất danh mục, có xu thế | 113 |
| 5.40 | Kiểm định ADF cho chuỗi lợi suất danh mục, có hệ số chặn | 114 |
| 5.41 | Kiểm định ADF cho chuỗi lợi suất danh mục, không có hệ số chặn | 115 |
| 5.42 | Mô hình ARIMA(1,0,1) cho chuỗi lợi suất danh mục | 115 |
| 5.43 | Mô hình ARIMA(1,0,0) cho chuỗi lợi suất danh mục | 116 |

Danh sách hình vẽ

| | | |
|------|--|----|
| 3.1 | Đồ thị doanh thu của VCB giai đoạn 2010-2024 | 5 |
| 3.2 | Biến động giá và lợi suất cổ phiếu Vietcombank giai đoạn 2023-2024 | 7 |
| 3.3 | Biểu đồ ACF chuỗi sai phân ΔVCB_t | 9 |
| 3.4 | Biểu đồ PACF chuỗi sai phân ΔVCB_t | 9 |
| 3.5 | Biểu đồ ACF chuỗi log - return VCB | 9 |
| 3.6 | Biểu đồ PACF chuỗi log - return VCB | 9 |
| 3.7 | Đồ thị doanh thu của CTG giai đoạn 2010-2024 | 11 |
| 3.8 | Biểu đồ chuỗi giá cổ phiếu và lợi suất của CTG | 13 |
| 3.9 | Đồ thị ACF và PACF của chuỗi ΔCTG_t | 14 |
| 3.10 | Đồ thị ACF và PACF của chuỗi log return | 15 |
| 3.11 | Đồ thị doanh thu của PVC giai đoạn 2010-2024 | 17 |
| 3.12 | Đồ thị chuỗi giá đóng cửa PVC giai đoạn 2023 - 2024 | 19 |
| 3.13 | Đồ thị chuỗi lợi suất PVC giai đoạn 2023 - 2024 | 19 |
| 3.14 | Biểu đồ ACF và PACF của chuỗi LogReturnPVC_t | 20 |
| 3.15 | Biểu đồ ACF và PACF của chuỗi ΔPVC_t | 21 |
| 3.16 | Kết quả mô hình Garch | 22 |
| 3.17 | Doanh thu của PVS giai đoạn 2010-2024 | 23 |
| 3.18 | Chuỗi giá cổ phiếu PVS | 25 |
| 3.19 | Chuỗi log-return của cổ phiếu PVS | 25 |
| 3.20 | ACF của ΔPVS_t | 26 |
| 3.21 | PACF của ΔPVS_t | 26 |
| 3.22 | ACF của log PVS_t | 27 |
| 3.23 | PACF của log PVS_t | 27 |
| 3.24 | Đồ thị biểu diễn u_t^2 | 28 |
| 3.25 | PACF của u_t^2 | 28 |
| 4.1 | Hệ số tương quan giữa các chuỗi giá | 29 |
| 4.2 | Hệ số tương quan giữa các chuỗi lợi suất | 29 |
| 4.3 | Hiệp phương sai giữa các chuỗi giá | 30 |
| 4.4 | Hiệp phương sai giữa các chuỗi lợi suất | 30 |
| 4.5 | Đồ thị chuỗi lợi suất danh mục giai đoạn 2023-2024 | 34 |
| 4.6 | Biểu đồ ACF chuỗi lợi suất danh mục | 35 |
| 4.7 | Biểu đồ PACF chuỗi lợi suất danh mục | 35 |
| 4.8 | Xác định bậc của trẽ | 38 |
| 4.9 | Kiểm định tương quan phần dư | 40 |
| 4.10 | Đồ thị dự báo r_{VCB} , r_{CTG} , r_{PVC} , r_{PVS} | 41 |
| 4.11 | Hàm phản ứng của r_{VCB} | 43 |
| 4.12 | Hàm phản ứng của r_{CTG} | 44 |
| 4.13 | Hàm phản ứng của r_{PVC} | 44 |
| 4.14 | Hàm phản ứng của r_{PVS} | 45 |
| 4.15 | Hàm phân ra phương sai của 4 chuỗi log-return | 46 |

| | | |
|-----|--|-----|
| 5.1 | Xem tính dừng qua nghiệm nghịch đảo | 66 |
| 5.2 | Kiểm định tính nhiễu trăng của phần dư | 67 |
| 5.3 | Xem tính dừng qua nghiệm nghịch đảo | 69 |
| 5.4 | Kiểm định tính nhiễu trăng của phần dư | 70 |
| 5.5 | Kết quả mô hình lin-lin | 84 |
| 5.6 | Xem tính dừng qua nghiệm nghịch đảo | 116 |
| 5.7 | Kiểm định tính nhiễu trăng của phần dư | 117 |

Chương 1

LỜI MỞ ĐẦU

Trong bối cảnh thị trường tài chính Việt Nam ngày càng phát triển và hội nhập sâu rộng với khu vực cũng như thế giới, việc phân tích và dự báo các chỉ số tài chính như doanh thu hay giá cổ phiếu ngày càng đóng vai trò quan trọng đối với nhà đầu tư, doanh nghiệp và nhà hoạch định chính sách. Đặc biệt, các ngành có vai trò trụ cột như ngân hàng và dầu khí thường có ảnh hưởng lớn đến diễn biến chung của thị trường.

Để có thể hiểu sâu hơn về tình hình kinh doanh và đưa ra những dự báo về triển vọng của hai ngành này, nhóm đã lựa chọn hai đại diện của mỗi ngành để phân tích. Cụ thể, trong ngành ngân hàng, nhóm chọn Ngân hàng TMCP Ngoại thương Việt Nam (VCB), Ngân hàng TMCP Công thương Việt Nam (CTG). Đối với ngành dầu khí, nhóm lựa chọn Tổng công ty Hóa chất và Dịch vụ Dầu khí - CTCP (PVC), và Tổng Công ty Cổ phần Dịch vụ Kỹ thuật Dầu khí Việt Nam (PVS).

Cụ thể, mục tiêu của nghiên cứu này là:

- Phân tích và dự báo doanh thu thuần của các ngân hàng và công ty này qua các quý.
- Áp dụng các kỹ thuật để dự báo giá cổ phiếu và mức độ rủi ro trong 10 phiên đầu năm 2025 của các ngân hàng và công ty này với dữ liệu trong 2 năm 2023 và 2024.
- Áp dụng các kỹ thuật tìm trọng số danh mục tối ưu dựa trên dữ liệu giá cổ phiếu, từ đó dự báo giá trị danh mục này trong 10 phiên đầu năm 2025
- Đưa ra các nhận xét và đánh giá về những sai sót trong dự báo.

Để đạt được mục tiêu nghiên cứu, nhóm đã áp dụng các phương pháp định lượng một cách hệ thống. Cụ thể, đối với chuỗi doanh thu, nhóm tiến hành thử nghiệm và so sánh các mô hình như lin-lin, lin-log, log-lin, log-log, cùng với các mô hình có yếu tố xu thế và mùa vụ, bao gồm cả mô hình Holt-Winters, nhằm xác định mô hình dự báo phù hợp nhất. Đối với giá cổ phiếu, nhóm sử dụng các mô hình ARIMA, VAR, GARCH và tiến hành so sánh kết quả dự báo giữa các mô hình này, qua đó đánh giá sai số dự báo giá cổ phiếu của các ngân hàng và công ty trong phạm vi nghiên cứu. Ngoài ra, nhóm cũng thực hiện xây dựng danh mục đầu tư tối ưu bằng cách xác định trọng số theo nhiều phương pháp khác nhau, từ đó đưa ra những nhận xét khách quan về độ chính xác của từng phương pháp dự báo.

Chương 2

TỔNG QUAN

2.1 Tổng quan các nhóm ngành nghiên cứu

2.1.1 Tổng quan ngành ngân hàng

Giai đoạn 2010–2024, ngành ngân hàng Việt Nam đã trải qua quá trình tái cấu trúc sâu rộng và chuyển mình mạnh mẽ. Trong giai đoạn 2010–2015, hệ thống đổi mới với khủng hoảng nợ xấu khi tỷ lệ nợ xấu từng lên tới xấp xỉ 17% nếu tính cả nợ tiềm ẩn, buộc Nhà nước phải mua lại 3 ngân hàng với giá 0 đồng (OceanBank, GPBank, VNCB) và thành lập VAMC vào năm 2013 để xử lý nợ xấu.

Từ 2016 đến 2019, ngành phục hồi rõ nét với tăng trưởng tín dụng trung bình khoảng 15%/năm, lợi nhuận của nhiều ngân hàng tăng mạnh, đặc biệt là khối tư nhân như Techcombank, MB, VPBank. Nợ xấu báo cáo giảm xuống dưới 3%, trong khi các dịch vụ ngân hàng điện tử và ngân hàng số bắt đầu phát triển nhanh chóng.

Giai đoạn 2020–2021, đại dịch COVID-19 khiến tăng trưởng tín dụng chững lại, còn khoảng 10–12%/năm, nhưng lại tạo cú hích mạnh mẽ cho chuyển đổi số, khi phần lớn giao dịch chuyển sang kênh trực tuyến. Ngân hàng Nhà nước cũng thực hiện nhiều chính sách hỗ trợ như cơ cấu nợ, giãn nợ và giảm lãi suất cho doanh nghiệp và người dân.

Từ 2022 đến 2024, ngành đổi mới như tín dụng tăng chậm (năm 2023 đạt khoảng 13,7%), thanh khoản căng thẳng và áp lực từ thị trường bất động sản trầm lắng. Tuy nhiên, hệ thống vẫn duy trì ổn định, nợ xấu được kiểm soát dưới 3%, nhiều ngân hàng hoàn thành chuẩn Basel II và bắt đầu triển khai Basel III.

Tính đến năm 2024, gần như tất cả ngân hàng lớn đã tích cực đầu tư vào nền tảng số, với hơn 95% giao dịch cá nhân thực hiện qua kênh số tại một số ngân hàng như MB, TPBank, Techcombank. Tỷ lệ bao phủ tín dụng cải thiện, hệ thống thanh toán không dùng tiền mặt mở rộng mạnh mẽ, và một số ngân hàng bắt đầu vươn ra thị trường quốc tế. Ngành ngân hàng Việt Nam đang hướng tới mô hình hiện đại, linh hoạt, đặt trọng tâm vào chuyển đổi số, quản trị rủi ro, và phát triển bền vững trong bối cảnh toàn cầu đầy biến động.

2.1.2 Tổng quan ngành dầu khí

Ngành dầu khí Việt Nam đã trải qua nhiều biến động và phát triển đáng kể trong giai đoạn 2010-2024. Đây là một trong những ngành kinh tế quan trọng của Việt Nam, đóng góp lớn vào ngân sách quốc gia và đảm bảo an ninh năng lượng.

Trong hơn 60 năm phát triển, ngành dầu khí đã xác định được trữ lượng dầu khí lớn và khai thác hiệu quả từ các mỏ trong nước và ngoài nước. Hiện nay, Việt Nam đang khai thác 25 mỏ dầu khí trong nước và 10 mỏ ngoài nước. Ngành dầu khí đã chuyển dịch từ một ngành chỉ khai thác và

xuất khẩu dầu thô sang một ngành có khả năng chế biến sâu, với các nhà máy lọc dầu như Dung Quất, góp phần cung cấp ổn định nhu cầu xăng dầu cho kinh tế và quốc phòng.

Giai đoạn 2010-2024, ngành dầu khí đối mặt với nhiều thách thức, bao gồm sự suy giảm tự nhiên của các mỏ dầu hiện hữu và xu hướng chuyển dịch năng lượng toàn cầu. Tuy nhiên, ngành đã chủ động thích nghi bằng cách đẩy mạnh thăm dò các đối tượng dầu khí phi truyền thống, phát triển công nghiệp khí, và tham gia vào lĩnh vực năng lượng tái tạo như điện gió ngoài khơi.

Năm 2024, ngành dầu khí đạt nhiều thành tựu ấn tượng. Sản lượng khai thác dầu đạt 9,87 triệu tấn, vượt 20,4% kế hoạch năm, và sản lượng khí đạt 6,32 tỷ m^3 , vượt 1,22 tỷ m^3 so với kế hoạch. Petrovietnam đạt tổng doanh thu vượt mốc 1 triệu tỷ đồng, nộp ngân sách gần 165.000 tỷ đồng, và có 3 phát hiện dầu khí mới trong năm.

Tổng quan, ngành dầu khí Việt Nam đã phát triển mạnh mẽ trong giai đoạn 2010-2024, đóng góp quan trọng vào nền kinh tế và đảm bảo an ninh năng lượng quốc gia. Mặc dù đối mặt với nhiều thách thức, ngành dầu khí đang chuyển mình để thích nghi với xu hướng chuyển dịch năng lượng toàn cầu.

2.2 Tổng quan về dữ liệu nghiên cứu

2.2.1 Thu thập số liệu

Nhóm sử dụng 499 quan sát về giá đóng cửa cổ phiếu của 4 mã cổ phiếu VCB, CTG (nhóm ngành ngân hàng) và PVS, PVC (nhóm ngành dầu khí). Dữ liệu được thu thập từ ngày 03/01/2023 đến ngày 31/12/2024.

Bên cạnh đó, dữ liệu theo quý bao gồm 60 quan sát về doanh thu, bắt đầu từ quý 1 năm 2010 đến quý 4 năm 2024 của cả 4 mã cổ phiếu.

Nguồn dữ liệu về doanh thu được thu thập từ báo cáo tài chính công bố trên trang FiinTrade, trong khi dữ liệu giá cổ phiếu được lấy từ trang Investing/ CaFeF.

2.2.2 Xử lý dữ liệu

Từ dữ liệu thu thập được, nhóm tính ra chuỗi log-return (chuỗi lợi suất liên tục) theo từng phiên của cả 4 mã cổ phiếu theo công thức sau:

$$r_{Y_t} = \ln\left(\frac{Y_t}{Y_{t-1}}\right) \times 100(%)$$

trong đó:

$$\begin{cases} r_t \text{ là lợi suất cổ phiếu phiên } t \\ Y_t \text{ là giá cổ phiếu phiên } t \\ Y_{t-1} \text{ là giá cổ phiếu phiên } t - 1 \end{cases}$$

Nhóm sử dụng phần mềm Rstudio và Excel để xử lý dữ liệu cũng như thực hiện các mô hình, kiểm định.

Chương 3

BÀI CÁ NHÂN

3.1 Phân tích chuỗi VCB - Phạm Hạnh Nguyên

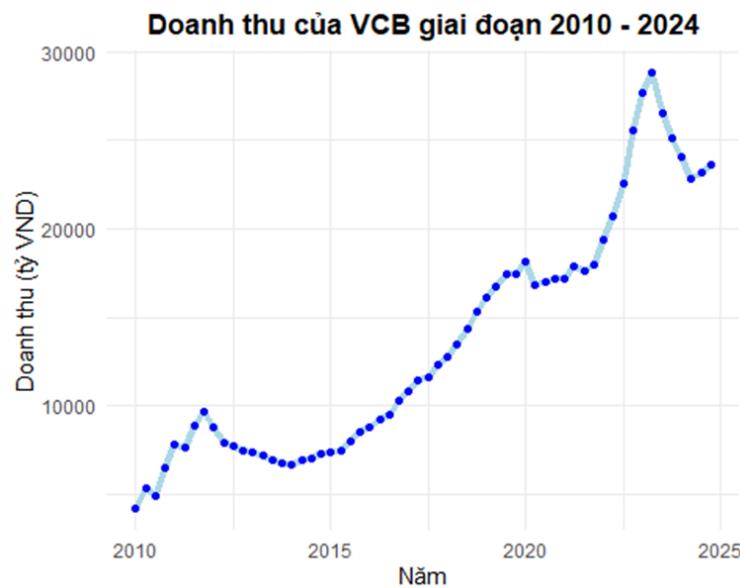
3.1.1 Tổng quan về Ngân hàng TMCP Ngoại thương Việt Nam

Ngân hàng Thương mại Cổ phần Ngoại thương Việt Nam (Vietcombank), tên giao dịch quốc tế là Joint Stock Commercial Bank for Foreign Trade of Vietnam, được thành lập vào ngày 1/4/1963, ban đầu là một ngân hàng chuyên doanh về ngoại thương trực thuộc Ngân hàng Nhà nước Việt Nam. Đây là một trong những doanh nghiệp niêm yết đầu tiên trên sàn HOSE (30/6/2009) và thường xuyên nằm trong nhóm VN30. Tính đến thời điểm gần đây, Vietcombank liên tục dẫn đầu về quy mô vốn hóa thị trường, hiệu quả hoạt động và mức độ uy tín trong ngành ngân hàng Việt Nam.

Dữ liệu được sử dụng trong bài nghiên cứu bao gồm 499 quan sát về giá cổ phiếu của mã VCB, được thu thập từ ngày 03/01/2023 đến ngày 31/12/2024. Từ đó thu được 498 giá trị log-return (lợi suất tính theo lãi suất liên tục) đã được tính toán nhằm phục vụ cho các phân tích định lượng. Bên cạnh đó, dữ liệu theo quý bao gồm 60 quan sát về doanh thu, bắt đầu từ quý 1 năm 2010 đến quý 4 năm 2024. Nguồn dữ liệu về doanh thu được thu thập từ báo cáo tài chính công bố trên trang FiinTrade, trong khi dữ liệu giá cổ phiếu được lấy từ trang Investing.

3.1.2 Dự báo chuỗi doanh thu năm 2025

Doanh thu của Vietcombank trong giai đoạn 2010–2024 cho thấy xu hướng tăng trưởng tích cực theo thời gian, mặc dù có một số giai đoạn biến động. Bắt đầu từ năm 2010, doanh thu nhìn chung có sự tăng trưởng nhưng không ổn định, cho tới năm 2012 thì giảm nhẹ và đi ngang trong vài năm tiếp theo. Từ năm 2015 đến 2019, doanh thu bắt đầu tăng trưởng rõ rệt và ổn định hơn, phản ánh sự mở rộng quy mô và hiệu quả hoạt động kinh doanh được cải thiện. Năm 2020 và 2021, tốc độ tăng trưởng doanh thu chững lại do COVID-19. Tuy nhiên bắt đầu từ 2022, doanh thu tăng trưởng mạnh mẽ, đạt đỉnh vào năm 2023 và sau đó tiếp tục giảm mạnh. Đến năm 2024, doanh thu đã có dấu hiệu phục hồi trở lại nhưng vẫn chưa đạt đỉnh trở lại như năm 2023. Nhìn chung, chuỗi doanh thu của Vietcombank phản ánh sự phát triển dài hạn bền vững, dù chịu tác động từ các yếu tố ngắn hạn của nền kinh tế.



Hình 3.1: Đồ thị doanh thu của VCB giai đoạn 2010-2024

Các mô hình dự báo doanh thu về bán hàng và cung cấp dịch vụ của Vietcombank với tập training set (2010 – 2023)

Ta sử dụng dữ liệu 56 quý, từ quý 1 năm 2010 đến quý 4 năm 2023 cho training set. Từ training set ta tạo ra các mô hình dự báo cho bốn quý năm 2024. So sánh dữ liệu được dự báo từ các mô hình với dữ liệu thực tế bằng RMSE và MAPE. Mô hình có RMSE và MAPE nhỏ nhất là mô hình

dự báo chính xác nhất, sẽ được sử dụng để dự báo cho năm 2025.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

Giai thích: cách tính ở mô hình (8) và mô hình (9):

- Mô hình (8): Holt - Winter có mùa vụ, dạng cộng:

$$DT_{T+k}^F = (25491.968 - 2113.555 \times k) + S_q$$

với $S_1 = 270.318, S_2 = -32.470, S_3 = -1209.929, S_4 = -394.177$

VD: Quý 3/2024: $DT_{56+3}^F = (25491.968 - 2113.555 \times 3) - 1209.929 = 17941.37$

- Mô hình (9): Holt - Winter có mùa vụ, dạng nhân:

$$DT_{T+k}^F = (26442.190 - 533.438 \times k)S_q$$

với $S_1 = 0.984, S_2 = 0.987, S_3 = 0.952, S_4 = 0.949$

VD: Quý 4/2024: $DT_{56+4}^F = (26442.190 - 533.438 \times 4) \times 0.949 = 23068.71$

| Quý 1 | Quý 2 | Quý 3 | Quý 4 | MAPE (%) | RMSE toàn mẫu | RMSE 2024 |
|---|----------|----------|----------|-------------------|---------------|-----------|
| 24019.66 | 22862.71 | 23191.21 | 23581.26 | (Giá trị thực tế) | | |
| (1) Mô hình lin - lin | | | | | | |
| $DT_t = 2471.94 + 363.05t + u_t$ | | | | | | |
| 23165.93 | 23528.98 | 23892.03 | 24255.09 | 19.476 | 2423.779 | 727.663 |
| (2) Mô hình lin - log | | | | | | |
| $DT_t = -3650.5 + 5351.2\ln(t) + u_t$ | | | | | | |
| 17984.53 | 18077.60 | 18169.07 | 18259.01 | 34.329 | 4207.721 | 5311.975 |
| (3) Mô hình log - lin | | | | | | |
| $\ln(DT_t) = 8.530198 + 0.028440t + u_t$ | | | | | | |
| 25624.18 | 26363.4 | 27123.95 | 27906.43 | 12.132 | 1709.233 | 3500.101 |
| (4) Mô hình log - log | | | | | | |
| $\ln(DT_t) = 7.94136 + 0.45468\ln(t) + u_t$ | | | | | | |
| 17670.62 | 17810.90 | 17949.88 | 18087.58 | 23.650 | 3710.075 | 5556.150 |
| (5) Hồi quy với biến giả mùa vụ | | | | | | |
| $DT_t = 12356.8 + 319.3s_2 + 506.9s_3 + 1022.3s_4 + u_t$ | | | | | | |
| 12356.79 | 12676.13 | 12863.72 | 13379.09 | 51.326 | 6338.129 | 10612.85 |
| (6) Xu thế tuyến tính + mùa vụ, dạng cộng | | | | | | |
| $DT_t = 2549.49 + 363.23t - 43.89s_2 - 219.54s_3 - 67.4s_4 + u_t$ | | | | | | |
| 23253.79 | 23573.13 | 23760.72 | 24276.09 | 19.444 | 2422.372 | 688.909 |
| (7) Xu thế tuyến tính + mùa vụ, dạng nhân | | | | | | |
| $DT_t = 2470.0313 + 363.4982t + 2.5235t \times s_2 - 3.4846t \times s_3 - 0.4262t \times s_4 + u_t$ | | | | | | |
| 23189.43 | 23699.29 | 23710.83 | 24254.35 | 19.498 | 2422.770 | 726.672 |
| (8) Holt - Winter có mùa vụ, dạng cộng | | | | | | |
| $DT_{T+k}^F = (25491.968 - 2113.555 \times k) + S_q$ | | | | | | |
| 23648.73 | 21232.39 | 17941.37 | 16643.57 | 6.475 | 1014.786 | 4429.669 |
| (9) Holt - Winter có mùa vụ, dạng nhân | | | | | | |
| $DT_{T+k}^F = (26442.190 + 533.438 \times k)S_q$ | | | | | | |
| 26548.40 | 27153.86 | 26707.81 | 27123.06 | 7.201 | 1127.600 | 3525.599 |

Bảng 3.1: Bảng kết quả dự báo cho chuỗi doanh thu VCB trong năm 2024

Bảng 3.1 cung cấp một số thông tin như giá trị doanh thu thực tế của Vietcombank năm 2024, tên mô hình dự báo, công thức, các giá trị dự báo trong bốn quý 2024 và các sai số. Lưu ý: dòng đầu tiên ngay dưới tiêu đề là giá trị thực tế của năm 2024.

Vì mô hình Holt - Winter mùa vụ dạng cộng có MAPE và RMSE toàn bộ mẫu nhỏ nhất nên ta sẽ dùng mô hình này để dự báo cho năm 2025.

Dự báo doanh thu năm 2025

Ước lượng mô hình Holt - Winter có mùa vụ dạng cộng cho tập dữ liệu giai đoạn 2010 - 2024, ta được kết quả như sau:

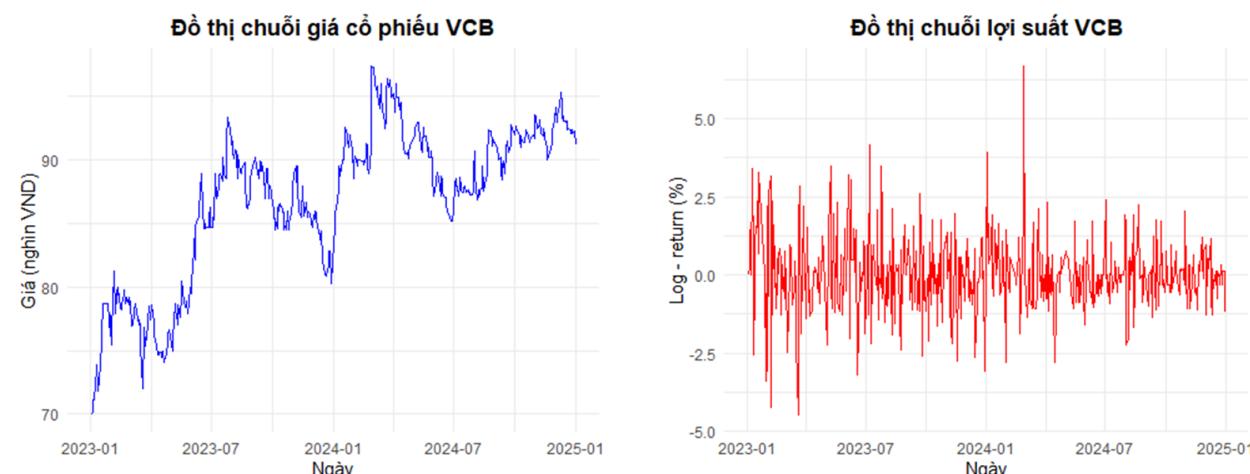
$$DT_{T+k}^F = (23716.907 + 89.995 \times k) + S_q$$

với $S_1 = 59.436, S_2 = 84.071, S_3 = -164.614, S_4 = -135.647$

| | Quý 1 | Quý 2 | Quý 3 | Quý 4 |
|-------------------------|----------|----------|----------|----------|
| Kết quả dự báo năm 2025 | 23866.34 | 23980.97 | 23822.28 | 23941.24 |

Bảng 3.2: Bảng kết quả dự báo cho chuỗi doanh thu VCB trong năm 2025

3.1.3 Dự báo giá cổ phiếu 10 phiên đầu năm 2025



(a) Đồ thị chuỗi giá cổ phiếu Vietcomabnk giai đoạn 2023-2024

(b) Đồ thị chuỗi lợi suất của Vietcomabnk giai đoạn 2023-2024

Hình 3.2: Biến động giá và lợi suất cổ phiếu Vietcombank giai đoạn 2023-2024

Nhìn chung, giá cổ phiếu VCB có xu hướng tăng nhẹ theo thời gian, bất chấp những biến động đáng kể trong ngắn hạn. Cụ thể, giá cổ phiếu khởi đầu ở mức xấp xỉ 70 nghìn VND và kết thúc chu kỳ quan sát ở mức trên 90 nghìn VND, cho thấy sự tăng trưởng tương đối tích cực. Tuy nhiên, từ 2023 tới tháng 7/2024, giá có nhiều đợt dao động mạnh với các đỉnh và đáy rõ rệt.

Lợi suất biến động mạnh trong năm 2023, đặc biệt là trong nửa đầu năm, với nhiều giá trị tăng và giảm đột biến vượt ngưỡng $\pm 2,5\%$. Một số điểm cực trị có thể vượt mức 5% hoặc dưới -3%, cho thấy những thời điểm có biến động giá cổ phiếu mạnh mẽ. Sang năm 2024, mức độ biến động có xu hướng giảm dần, lợi suất trở nên ổn định hơn, dao động chủ yếu quanh mức 0%.

Dự báo 10 phiên cổ phiếu đầu tiên năm 2025

- Bước 1: Kiểm định tính dừng:

| Cập giả thiết | Chuỗi giá VCB_t | Chuỗi sai phân ΔVCB_t | Chuỗi log-return VCB_t |
|--|--|--|--|
| Kiểm định ADF trend: H_0 : Chuỗi có nghiệm đơn vị (không dừng) H_1 : Chuỗi không có nghiệm đơn vị (dừng xu thế) | $ \tau_{qs} = 3.3261 < \tau_{0.05} = 3.42$ → Chưa bác bỏ H_0 , chuỗi VCB_t không dừng. $p\text{-value} = 0.0572 > 0.05$ → Hệ số xu thế tt không có ý nghĩa thống kê. | $ \tau_{qs} = 16.8717 > \tau_{0.05} = 3.42$ → Bác bỏ H_0 , chuỗi ΔVCB_t dừng. $p\text{-value} = 0.332 > 0.05$ → Hệ số xu thế tt không có ý nghĩa thống kê. | $ \tau_{qs} = 16.9636 > \tau_{0.05} = 3.42$ → Bác bỏ H_0 , chuỗi log - return VCB_t dừng. $p\text{-value} = 0.285 > 0.05$ → Hệ số xu thế tt không có ý nghĩa thống kê. |
| Kiểm định ADF drift: H_0 : Chuỗi có nghiệm đơn vị (không dừng) H_1 : Chuỗi không có nghiệm đơn vị (dừng vây quanh trung bình dài hạn) | $ \tau_{qs} = 2.8982 > \tau_{0.05} = 2.87$ → Bác bỏ H_0 , chuỗi VCB_t dừng. $p\text{-value} = 0.00325 < 0.05$ → Hệ số chặn có ý nghĩa thống kê. | $ \tau_{qs} = 16.8448 > \tau_{0.05} = 2.87$ → Bác bỏ H_0 , chuỗi ΔVCB_t dừng. $p\text{-value} = 0.330 > 0.05$ → Hệ số chặn không có ý nghĩa thống kê. | $ \tau_{qs} = 16.9274 > \tau_{0.05} = 2.87$ → Bác bỏ H_0 , chuỗi log - return VCB_t dừng. $p\text{-value} = 0.301 > 0.05$ → Hệ số chặn không có ý nghĩa thống kê. |
| Kiểm định ADF none: H_0 : Chuỗi có nghiệm đơn vị (không dừng) H_1 : Chuỗi không có nghiệm đơn vị (dừng vây quanh 0) | $ \tau_{qs} = 0.7747 < \tau_{0.05} = 1.95$ → Chưa bác bỏ H_0 , chuỗi VCB_t không dừng. $\tau_{qs} = 0.7747 > 0$ Nên thêm yếu tố hệ số chặn hoặc xu thế. | $ \tau_{qs} = 16.8177 > \tau_{0.05} = 1.95$ → Bác bỏ H_0 , chuỗi ΔVCB_t dừng. $\tau_{qs} = -16.8177 < 0$ Không cần thêm yếu tố hệ số chặn hoặc xu thế. | $ \tau_{qs} = 16.8948 > \tau_{0.05} = 1.95$ → Bác bỏ H_0 , chuỗi log - return VCB_t dừng. $\tau_{qs} = -16.8948 < 0$ Không cần thêm yếu tố hệ số chặn hoặc xu thế. |

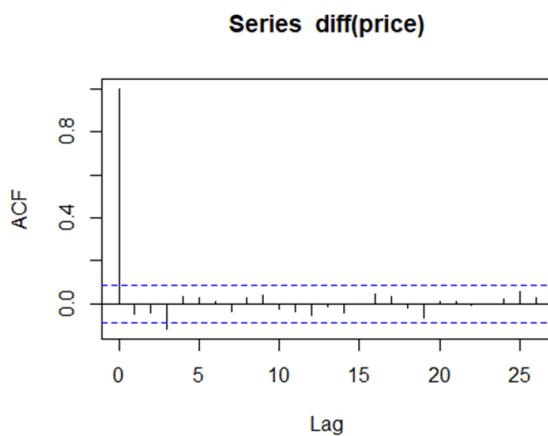
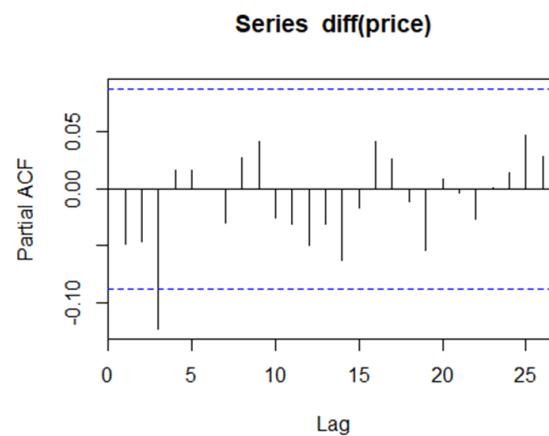
Bảng 3.3: Bảng kết quả kiểm định tính dừng chuỗi VCB ở mức ý nghĩa 5%

Kết luận: Từ kết quả trên, ta nhận thấy chuỗi giá VCB dừng sai phân và không có hệ số chặn và chuỗi log - return VCB là chuỗi dừng vây quanh 0.

- Bước 2: Chọn mô hình ARIMA phù hợp để ước lượng, kiểm tra, đánh giá và dự báo:

+ Cho chuỗi sai phân ΔVCB_t :

Từ 2 đồ thị trên, có thể thấy tương quan giữa biến ΔVCB_t với trễ tại bậc 3 và tương quan với shock của 3 kì trước đó. Ta đề xuất mô hình ARIMA(3,1,3), ARIMA(1,1,3), ARIMA(3,1,1), ARIMA(3,1,0)

Hình 3.3: Biểu đồ ACF chuỗi sai phân ΔVCB_t Hình 3.4: Biểu đồ PACF chuỗi sai phân ΔVCB_t

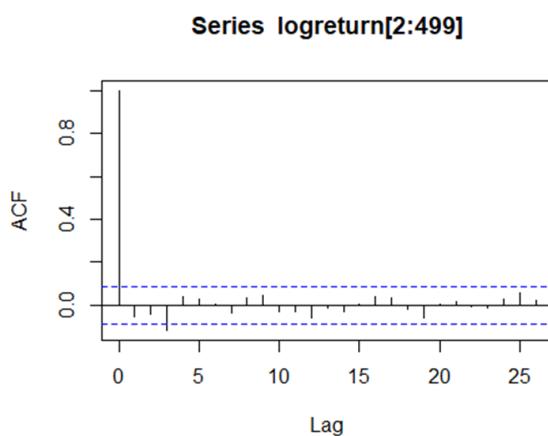
| Thông số | ARIMA(3,1,3) | ARIMA(1,1,3) | ARIMA(3,1,1) | ARIMA(3,1,0) |
|-------------------------|--------------|--------------|--------------|----------------|
| Hệ số | | | | |
| ar1 | 0.3957 | -0.1812 | -0.1721 | -0.0545 |
| ar2 | -0.8782* | — | -0.0560 | -0.0503 |
| ar3 | 0.3403 | — | -0.1259* | -0.1204* |
| ma1 | -0.4475 | 0.1294 | — | — |
| ma2 | 0.8461* | -0.0506 | — | — |
| ma3 | -0.4446 | -0.1169* | 0.1194 | — |
| Chỉ tiêu mô hình | | | | |
| AIC | 1437.55 | 1436.41 | 1435.83 | 1433.97 |

Bảng 3.4: Bảng so sánh các mô hình ARIMA cho chuỗi sai phân ΔVCB_t

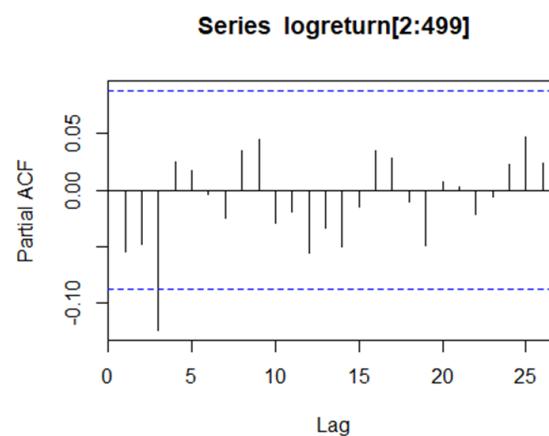
Ghi chú: * $p < 0.05$, hệ số có ý nghĩa thống kê.

Mô hình ARIMA(3,1,0) có hệ số AIC thấp nhất, ta chọn mô hình này để dự báo chuỗi giá VCB.

+ Cho chuỗi log - return VCB:



Hình 3.5: Biểu đồ ACF chuỗi log - return VCB



Hình 3.6: Biểu đồ PACF chuỗi log - return VCB

Tương tự với biểu đồ ACF và PACF của chuỗi sai phân ΔVCB_t , tương quan giữa log - return VCB_t với trễ tại bậc 3 và tương quan với shock của 3 kì trước đó. Đề xuất mô hình ARIMA(3,0,3),

ARIMA(1,0,3), ARIMA(3,0,1), ARIMA(3,0,0)

| Thông số | ARIMA(3,0,3) | ARIMA(1,0,3) | ARIMA(3,0,1) | ARIMA(3,0,0) |
|-------------------------|--------------|--------------|--------------|----------------|
| Hệ số | | | | |
| ar1 | 0.0718 | -0.2308 | -0.2157 | -0.0631 |
| ar2 | -0.6728 | – | -0.0637 | -0.0551 |
| ar3 | 0.0393 | – | -0.1312* | -0.1235* |
| ma1 | -0.1327 | 0.1714 | – | – |
| ma2 | 0.6315 | -0.0583 | – | – |
| ma3 | -0.1701* | -0.1218* | 0.1551 | – |
| Chỉ tiêu mô hình | | | | |
| AIC | 1600.57 | 1597.88 | 1597.27 | 1595.51 |

Bảng 3.5: Bảng so sánh các mô hình ARIMA cho chuỗi log - return VCB

Ghi chú: * $p < 0.05$, hệ số có ý nghĩa thống kê.

Mô hình ARIMA(3,0,0) có hệ số AIC thấp nhất, vậy nên ta chọn mô hình này để dự báo chuỗi giá VCB_t .

- Bước 3: Kiểm định về mô hình: Kết quả phân tích được trình bày dưới phần phụ lục cho thấy rằng tất cả các mô hình ARIMA sử dụng đều có nghiệm nghịch đảo nằm trong đường tròn đơn vị không có tự tương quan giữa các bậc, hay phần dư là nhiễu trắng.

- Bước 4: Dự báo cho 10 phiên đầu năm 2025: Từ các mô hình trên, ta có được kết quả dự báo chuỗi giá đóng cửa VCB cho 10 phiên đầu năm 2025. Với mô hình ARIMA cho chuỗi log - return, ta suy ra chuỗi giá theo công thức sau để so sánh với chuỗi giá thực:

$$VCB_t = VCB_{t-1} \times e^{\frac{\log VCB_t}{100}}$$

| Phiên | Giá dự báo (nghìn VND) | Log - return dự báo (%) | Giá quy đổi từ log - return (nghìn VND) | Giá thực (nghìn VND) |
|-------------|---------------------------|----------------------------|---|-------------------------|
| 1 | 91.243 | 0.1224 | 91.312 | 91.900 |
| 2 | 91.284 | 0.1111 | 91.413 | 92.000 |
| 3 | 91.412 | 0.2005 | 91.597 | 92.900 |
| 4 | 91.398 | 0.0323 | 91.626 | 92.300 |
| 5 | 91.387 | 0.0394 | 91.662 | 92.400 |
| 6 | 91.373 | 0.0371 | 91.696 | 91.900 |
| 7 | 91.376 | 0.0577 | 91.749 | 91.400 |
| 8 | 91.378 | 0.0556 | 91.800 | 91.900 |
| 9 | 91.379 | 0.0549 | 91.851 | 91.700 |
| 10 | 91.379 | 0.0525 | 91.899 | 91.400 |
| RMSE | 751.8929 | | 619.0739 | |
| MAPE | 0.6777% | | 0.5632% | |

Bảng 3.6: Dự báo giá cổ phiếu VCB 10 phiên đầu năm 2025

Từ bảng 4.17 ta thấy cách dự báo từ chuỗi log - return VCB là tối ưu nhất với sai số dự báo RMSE và MAPE của cách này nhỏ hơn so với dự báo trực tiếp chuỗi giá đóng cửa. Do đó, ta sử dụng mô hình ARIMA(3,0,0) cho chuỗi log - return VCB để dự báo cho 10 phiên đầu năm 2025. Ta cũng sẽ dùng kết quả dự báo từ chuỗi log - return VCB để làm cho những bài tập nhóm.

3.2 Phân tích chuỗi CTG - Trần Thị Kiều Oanh

3.2.1 Tổng quan về ngân hàng Viettinbank

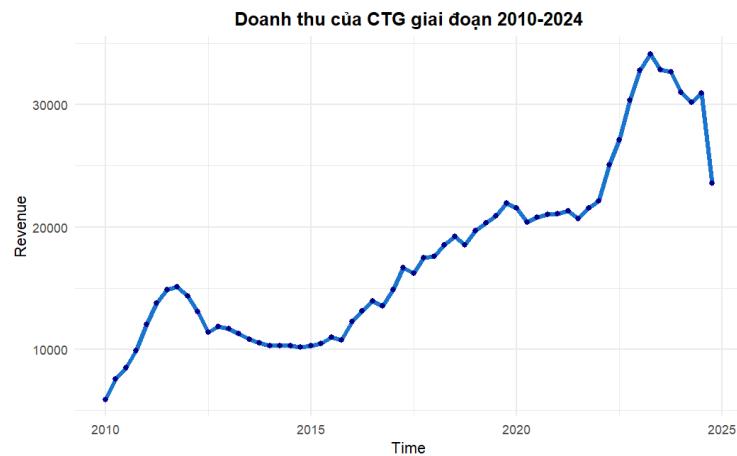
Ngân hàng TMCP Công thương Việt Nam (VietinBank) – một trong những ngân hàng thương mại lớn, giữ vai trò quan trọng, trụ cột của ngành ngân hàng Việt Nam. VietinBank có tiền thân là Ngân hàng Công thương Việt Nam, được thành lập dưới tên gọi Ngân hàng chuyên doanh Công thương Việt Nam ngày 26/03/1988, chính thức được đổi tên thành “Ngân hàng Công thương Việt Nam” ngày 14/11/1990.

Ngày 25/12/2008, Ngân hàng Công thương Việt Nam tổ chức bán đấu giá cổ phần ra công chúng thành công và thực hiện chuyển đổi thành doanh nghiệp cổ phần. Ngân hàng Thương mại Cổ phần Công thương Việt Nam chính thức hoạt động theo giấy chứng nhận đăng ký kinh doanh lần đầu do Sở Kế hoạch và Đầu tư TP Hà Nội cấp ngày 03/07/2009. Ngày 16/07/2009, cổ phiếu của VietinBank chính thức được niêm yết và giao dịch trên HOSE với mã giao dịch là **CTG**.

Dữ liệu được sử dụng trong bài nghiên cứu bao gồm 499 quan sát về giá cổ phiếu của mã CTG, được thu thập từ ngày 03/01/2023 đến ngày 31/12/2024. Từ các mức giá này, 498 giá trị log-return (lợi suất tính theo lãi suất liên tục) đã được tính toán nhằm phục vụ cho các phân tích định lượng. Bên cạnh đó, dữ liệu theo quý bao gồm 60 quan sát về doanh thu, bắt đầu từ quý I năm 2010 đến quý IV năm 2024. Nguồn dữ liệu về doanh thu được thu thập từ báo cáo tài chính công bố trên trang *VietstockFinance*, trong khi dữ liệu giá cổ phiếu được lấy từ trang *CafeF*.

3.2.2 Dự báo chuỗi doanh thu năm 2025

Từ quý I/2010 đến quý IV/2024, doanh thu của VietinBank nhìn chung tăng ổn định, dù có biến động trong một số giai đoạn. Bắt đầu ở mức 5.905 tỷ đồng (Q1/2010), doanh thu tăng mạnh, đạt 15.126 tỷ đồng (Q4/2011), đặc biệt Q1/2011 tăng 104,15 so với cùng kỳ. Tuy nhiên, từ 2012–2015, doanh thu giảm dần, xuống còn 10.276 tỷ đồng (Q1/2015), phản ánh giai đoạn tăng trưởng chậm lại. Từ 2016, doanh thu phục hồi ổn định, đạt 16.681 tỷ đồng vào Q4/2017. Năm 2020, dịch Covid-19 khiến doanh thu chững lại. Từ 2021, doanh thu phục hồi mạnh, đạt đỉnh 34.116 tỷ đồng (Q2/2023) trước khi giảm xuống 23.581262 tỷ đồng (Q4/2024).



Hình 3.7: Đồ thị doanh thu của CTG giai đoạn 2010-2024

Các mô hình dự báo doanh thu thuần về bán hàng và cung cấp dịch vụ của Viettinbank với tập training set (2010 – 2023)

Ta sử dụng dữ liệu 56 quý, từ quý 1 năm 2010 đến quý 4 năm 2023 cho training set. Từ training set ta tạo ra các mô hình dự báo cho bốn quý năm 2024. So sánh dữ liệu được dự báo từ các mô hình với dữ liệu thực tế bằng RMSE và MAPE.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

Mô hình có RMSE và MAPE nhỏ nhất là mô hình dự báo chính xác nhất, sẽ được sử dụng để dự báo cho năm 2025.

Dưới đây ta có bảng dự báo kết quả cho chuỗi doanh thu của Viettinbank trong năm 2024.

Lưu ý: dòng đầu là giá trị thực tế của năm 2024, từ dòng sau là tên mô hình, công thức và các giá trị dự báo.

| Quý 1 | Quý 2 | Quý 3 | Quý 4 | MAPE (%) | RMSE toàn mẫu | RMSE 2024 |
|---|-----------|-----------|-----------|-------------------|---------------|-----------|
| 30951.965 | 30138.279 | 30924.362 | 23581.262 | (Giá trị thực tế) | | |
| (1) Mô hình lin - lin | | | | | | |
| $DT_t = 6166.60 + 376.54t + u_t$ | | | | | | |
| 27629.38 | 28005.92 | 28382.46 | 28759.00 | 16.186 | 3104.764 | 3494.88 |
| (2) Mô hình lin - log | | | | | | |
| $DT_t = 336.8 + 5599.8ln(t) + u_t$ | | | | | | |
| 21717.9 | 21884.5 | 22050.3 | 22216.0 | 25.504 | 4682.524 | 5311.98 |
| (3) Mô hình log - lin | | | | | | |
| $ln(DT_t) = 9.028276 + 0.022050t + u_t$ | | | | | | |
| 29212.17 | 29860.44 | 30523.47 | 31201.79 | 13.12 | 2549.862 | 3922.53 |
| (4) Mô hình log - log | | | | | | |
| $ln(DT_t) = 8.55845 + 0.35684ln(t) + u_t$ | | | | | | |
| 22024.05 | 22159.92 | 22295.82 | 22431.74 | 19.338 | 4279.057 | 5556.150 |
| (5) Hồi quy với biến giả mùa vụ | | | | | | |
| $DT_t = 16175.7 + 676.9s_2 + 858.7s_3 + 1353.6s_4 + u_t$ | | | | | | |
| 16175.7 | 16852.6 | 17034.4 | 17529.3 | 38.07118 | 6815.061 | 10612.85 |
| (6) Xu thế tuyến tính + mùa vụ, dạng cộng | | | | | | |
| $DT_t = 6015.32 + 376.31t + 300.57s_2 + 106.03s_3 + 224.68s_4 + u_t$ | | | | | | |
| 27559.99 | 28236.87 | 28656.67 | 29138.9 | 16.142 | 3012.654 | 3610.59 |
| (7) Xu thế tuyến tính + mùa vụ, dạng nhân | | | | | | |
| $DT_t = 6174.995 + 368.514t + 11.527ts_2 + 6.296ts_3 + 12.536ts_4 + u_t$ | | | | | | |
| 27121.3 | 27953.6 | 28452.3 | 28996.1 | 16.17 | 3100.61 | 3711.10 |
| (8) Xu thế không tuyến tính + mùa vụ, dạng cộng | | | | | | |
| $ln(DT_t) = 9.013454 + 0.022024t + 0.025128s_2 + 0.016236s_3 + 0.020847s_4 + u_t$ | | | | | | |
| 27194 | 28047 | 28699 | 29483 | 13.279 | 2548.503 | 3645.32 |
| (9) Xu thế không tuyến tính + mùa vụ, dạng nhân | | | | | | |
| $ln(DT_t) = 9.02834 + 0.0218.t + 0.00042.t.s_2 + 0.0001216.t.s_3 + 0.0001985.t.s_4 + u_t$ | | | | | | |
| 29009.47 | 29880.65 | 30776.77 | 31700.32 | 13.28 | 2546.72 | 3685.56 |
| (10) Holt - Winter có mùa vụ, dạng cộng | | | | | | |
| $DT_{T+k}^F = (33027.0349 + 261.3418k) + S_q$ | | | | | | |
| 33481.19 | 33996.33 | 33701.42 | 33603.61 | 5.679 | 1171.123 | 4429.669 |
| (11) Holt - Winter có mùa vụ, dạng nhân | | | | | | |
| $DT_{T+k}^F = (33710.11 + 865.10k)S_q$ | | | | | | |
| 34930.45 | 36187.17 | 35821.17 | 35406.9 | 6.035 | 1297.908 | 3525.6 |

Bảng 3.7: Bảng kết quả dự báo cho chuỗi doanh thu CTG trong năm 2024

Giải thích: cách tính ở mô hình (10) và mô hình (11):

- Mô hình (10): Holt - Winter có mùa vụ, dạng cộng:

$$DT_{T+k}^F = (33027.0349 + 261.k)3418 + S_q$$

với $S_1 = 192.8182, S_2 = 446.6110, S_3 = -249.6323, S_4 = -389.7969$

VD: Quý 3/2024: $DT_{56+3}^F = (33027.0349 + 261.3418 \times 3) - 249.6323 = 33561.43$

- Mô hình (11): Holt - Winter có mùa vụ, dạng nhân:

$$DT_{T+k}^F = (33710.11 + 865.10k)S_q$$

với $S_1 = 1.005, S_2 = 1.032, S_3 = 0.989, S_4 = 0.968$

VD: Quý 4/2024: $DT_{56+4}^F = (3.371011e + 04 - 8.651010e + 02 \times 4) \times 0.968 = 29281.72$

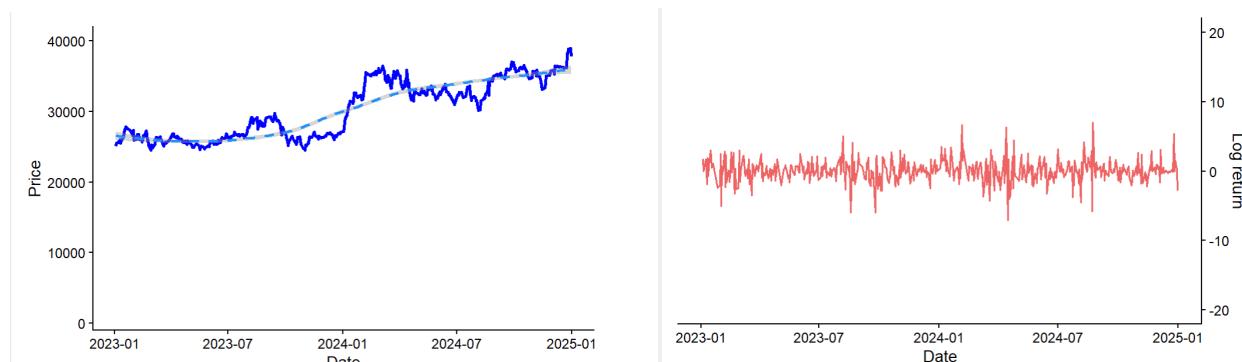
Dự báo doanh thu năm 2025

Vì mô hình Holt - Winter có mùa vụ dạng cộng có MAPE và RMSE toàn bộ mẫu nhỏ nhất nên ta sẽ dùng mô hình này để dự báo cho năm 2025.

| | Quý 1 | Quý 2 | Quý 3 | Quý 4 |
|-------------------------|----------|----------|----------|----------|
| Kết quả dự báo năm 2025 | 34847.35 | 35431.39 | 36270.39 | 34666.93 |

Bảng 3.8: Bảng kết quả dự báo cho chuỗi doanh thu CTG trong năm 2025

3.2.3 Dự báo giá cổ phiếu 10 phiên đầu năm 2025



(a) Đồ thị chuỗi giá cổ phiếu Viettinbank giai đoạn 2023-2024

(b) Đồ thị chuỗi lợi suất của Viettinbank giai đoạn 2023-2024

Hình 3.8: Biểu đồ chuỗi giá cổ phiếu và lợi suất của CTG

Dựa trên hình (a), chuỗi giá cổ phiếu CTG giai đoạn 2023–2024 thể hiện rõ tính không dừng với xu hướng tăng đều theo thời gian — đặc điểm phổ biến của chuỗi giá tài chính. Ngược lại, hình (b) cho thấy chuỗi log-return dao động quanh giá trị trung bình gần 0, không có xu hướng rõ rệt, cho thấy đây có thể là chuỗi dừng. Điều này phù hợp với lý thuyết tài chính rằng lợi suất ổn định hơn giá. Tuy nhiên, cần kiểm định ADF để xác định chính xác tính dừng của chuỗi log-return.

Mô hình dự báo giá 10 phiên cổ phiếu đầu tiên năm 2025

Bước 1: Kiểm định tính dừng của chuỗi giá cổ phiếu CTG

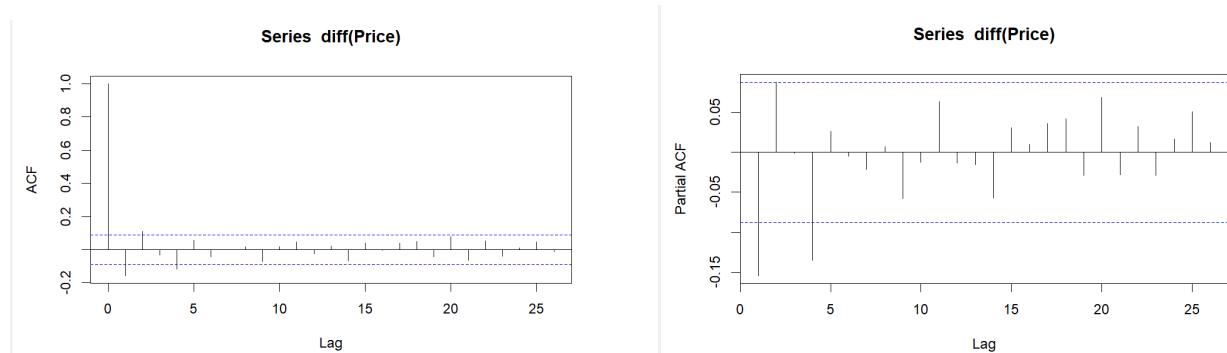
| Cặp giả thiết | Chuỗi giá CTG_t | Chuỗi sai phân ΔCTG_t | Chuỗi log-return CTG_t |
|--|---|---|---|
| Kiểm định ADF trend: H_0 : Chuỗi có nghiệm đơn vị (không dừng) H_1 : Chuỗi không có nghiệm đơn vị (dừng xu thế) | $ \tau_{qs} = 2.8933 < \tau_{0.05} = 3.96$ → Chưa bác bỏ H_0 , chuỗi CTG_t không dừng. $p\text{-value} = 0.00276 < 0.05$ → Hệ số xu thế tt có ý nghĩa thống kê. | $ \tau_{qs} = 15.36 > \tau_{0.05} = 3.96$ → Bác bỏ H_0 , chuỗi ΔCTG_t dừng. $p\text{-value} = 0.8789 > 0.05$ → Hệ số xu thế tt không có ý nghĩa thống kê. | $ \tau_{qs} = 15.49 > \tau_{0.05} = 3.96$ → Bác bỏ H_0 , chuỗi log - return CTG_t dừng. $p\text{-value} = 0.39 > 0.05$ → Hệ số xu thế tt không có ý nghĩa thống kê. |
| Kiểm định ADF drift: H_0 : Chuỗi có nghiệm đơn vị (không dừng) H_1 : Chuỗi không có nghiệm đơn vị (dừng vây quanh trung bình dài hạn) | $ \tau_{qs} = 1.2186 < \tau_{0.05} = 3.43$ → Chưa bác bỏ H_0 , chuỗi CTG_t không dừng. $p\text{-value} = 0.178 > 0.05$ → Hệ số chặn không có ý nghĩa thống kê. | $ \tau_{qs} = 15.366 > \tau_{0.05} = 3.43$ → Bác bỏ H_0 , chuỗi ΔCTG_t dừng. $p\text{-value} = 0.243 > 0.05$ → Hệ số chặn không có ý nghĩa thống kê. | $ \tau_{qs} = 15.509 > \tau_{0.05} = 3.43$ → Bác bỏ H_0 , chuỗi log - return CTG_t dừng. $p\text{-value} = 0.256 > 0.05$ → Hệ số chặn không có ý nghĩa thống kê. |
| Kiểm định ADF none: H_0 : Chuỗi có nghiệm đơn vị (không dừng) H_1 : Chuỗi không có nghiệm đơn vị (dừng vây quanh 0) | $ \tau_{qs} = 0.8595 < \tau_{0.05} = 2.58$ → Chưa bác bỏ H_0 , chuỗi CTG_t dừng. $\tau_{qs} = 0.8595 > 0$ Nên thêm yếu tố hệ số chặn hoặc xu thế. | $ \tau_{qs} = 15.4997 > \tau_{0.05} = 2.58$ → Bác bỏ H_0 , chuỗi ΔCTG_t dừng. $\tau_{qs} = -15.4997 < 0$ Không cần thêm yếu tố hệ số chặn hoặc xu thế. | $ \tau_{qs} = 15.317 > \tau_{0.05} = 2.58$ → Bác bỏ H_0 , chuỗi log - return CTG_t dừng. $\tau_{qs} = -15.317 < 0$ Không cần thêm yếu tố hệ số chặn hoặc xu thế. |

Bảng 3.9: Bảng kết quả kiểm định tính dừng chuỗi CTG ở mức ý nghĩa 5%

Nhận xét : Từ kết quả trên, thấy chuỗi giá CTG dừng sai phân và không có hệ số chặn ; chuỗi log return CTG là chuỗi dừng vây quanh 0

Bước 2: Chọn mô hình ARIMA phù hợp để ước lượng, kiểm tra, đánh giá và dự báo

+ Cho chuỗi sai phân Để xác định bậc $ARMA(p, q)$ phù hợp cho chuỗi ΔCTG_t , ta cần phân tích đồ thị hàm tự tương quan (ACF) và hàm tự tương quan riêng phần (PACF) của chuỗi này.



Hình 3.9: Đồ thị ACF và PACF của chuỗi ΔCTG_t

Dựa trên hai đồ thị ACF và PACF của chuỗi sai phân ΔCTG_t , ta đề xuất một số mô hình tiêu biểu để ước lượng chuỗi giá đóng cửa CTG_t là ARIMA(1,1,1), ARIMA(4,1,4), ARIMA(1,1,4), ARIMA(4,1,1)

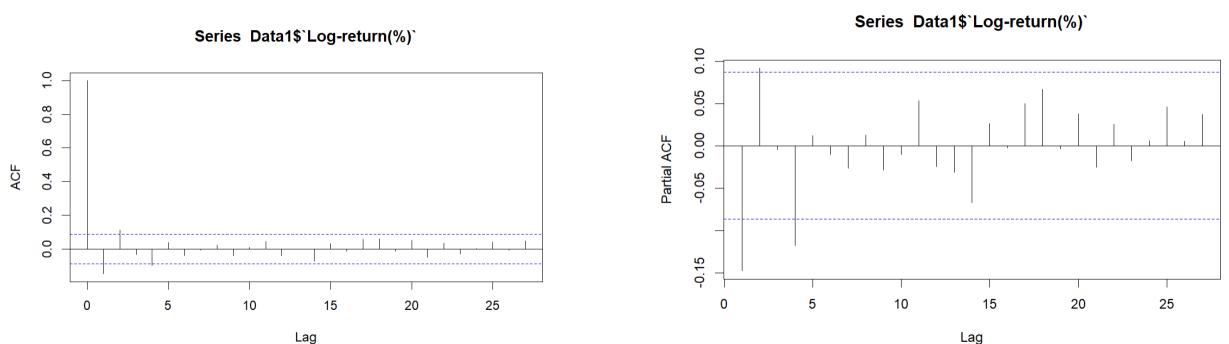
| Thông số | ARIMA(1,1,1) | ARIMA(4,1,4) | ARIMA(1,1,4) | ARIMA(4,1,1) |
|-------------------------|--------------|---------------|--------------|----------------|
| Hệ số | | | | |
| ar1 | -0.4373** | -0.1089 | -0.5325 | -0.3368 |
| ar2 | — | 0.2357 | — | 0.0746 |
| ar3 | — | 0.3030 | — | -0.0020 |
| ar4 | — | -0.1958 | — | -0.1380 |
| ma1 | 0.2821 | -0.0276 | 0.3926 | 0.1981 |
| ma2 | — | -0.1328 | 0.0428 | — |
| ma3 | — | -0.3163 | 0.0078 | — |
| ma4 | — | 0.1018 | -0.1282 | — |
| drift | 25.7488 | 25.3971 | 25.6306 | 25.6190 |
| Chỉ tiêu mô hình | | | | |
| AIC | 7601.15 | 7600.57 | 7596.59 | 7595.72 |
| RMSE | 494.56 | 488.30 | 489.31 | 488.88 |
| MAPE (%) | 1.14 | 1.14 | 1.14 | 1.14 |

Bảng 3.10: So sánh các mô hình ARIMA cho chuỗi giá đóng cửa CTG

Ghi chú: ** p < 0.05; * p < 0.1. Giá trị AIC càng nhỏ càng tốt. RMSE và MAPE càng nhỏ càng tốt.

Trong các mô hình đã ước lượng, **ARIMA(1,1,1)** có các hệ số ar1 và ma1 có ý nghĩa thống kê, trong khi **ARIMA(4,1,4)** và **ARIMA(1,1,4)** thiếu ý nghĩa rõ ràng. Mô hình **ARIMA(4,1,1)** dù hệ số không nổi bật nhưng có AIC thấp nhất (7595.72) và sai số dự báo tốt. Vì vậy, **ARIMA(4,1,1)** được chọn là mô hình phù hợp nhất để dự báo giá cổ phiếu **CTG**.

+ Đối với chuỗi log return



Hình 3.10: Đồ thị ACF và PACF của chuỗi log return

Dựa trên hai đồ thị ACF và PACF của chuỗi log return CTG, ta đề xuất một số mô hình tiêu biểu để ước lượng chuỗi log return CTG_t là ARIMA(0,0,4), ARIMA(2,0,2)

| Chỉ tiêu | ARIMA(0,0,4) | ARIMA(2,0,2) |
|----------|---------------|---------------|
| Hệ số | MA1 = -0.132 | AR1 = -1.0000 |
| | MA2 = 0.1158 | AR2 = -0.6546 |
| | MA3 = -0.0580 | MA1 = 0.8678 |
| | MA4 = -0.0847 | MA2 = 0.6462 |
| | Mean = 0.0821 | Mean = 0.0826 |
| AIC | 1886.78 | 1886.46 |
| RMSE | 1.5893 | 1.5887 |
| MAPE | Inf | Inf |

Bảng 3.11: So sánh mô hình ARIMA(0,0,4) và ARIMA(2,0,2)

Nhận xét: Dựa trên bảng so sánh, hai mô hình ARIMA(0,0,4) và ARIMA(2,0,2) có mức độ phù hợp gần như tương đương, thể hiện qua chỉ số AIC và RMSE chênh lệch không đáng kể. Tuy nhiên, do cả hai đều có MAPE bằng vô cùng, nên không thể so sánh chính xác về sai số phần trăm. Trong khi ARIMA(2,0,2) có cấu trúc phức tạp hơn với hệ số AR lớn, ARIMA(0,0,4) lại đơn giản và ổn định hơn. Do đó, ARIMA(0,0,4) được xem là lựa chọn hợp lý nếu ưu tiên mô hình gọn nhẹ và ít nguy cơ overfitting.

Bước 3 : Kiểm định về mô hình

Kết quả phân tích được trình bày dưới phần phụ lục cho thấy rằng tất cả các mô hình ARIMA sử dụng đều có nghiêm nghịch đảo nằm trong đường tròn đơn vị không có tự tương quan giữa các bậc, hay phần dư là nhiễu trắng.

Bước 4: Dự báo log return cho 10 phiên đầu năm 2025

| Phiên | Chuỗi giá CTG | | | | Chuỗi log-return CTG | | Giá thực |
|-------|---------------|--------------|--------------|--------------|----------------------|--------------|----------|
| | ARIMA(1,1,1) | ARIMA(4,1,4) | ARIMA(1,1,4) | ARIMA(4,1,1) | ARIMA(0,0,4) | ARIMA(2,1,2) | |
| 1 | 37994.58 | 38050.10 | 38038.90 | 38032.77 | 37184.41 | 37112.77 | 38100.00 |
| 2 | 37946.49 | 37795.28 | 37862.72 | 37825.25 | 37184.68 | 37149.47 | 37200.00 |
| 3 | 38004.53 | 37859.41 | 37986.96 | 37956.36 | 37297.06 | 37241.28 | 37200.00 |
| 4 | 38016.16 | 38009.32 | 38068.60 | 38071.03 | 37358.98 | 37200.66 | 37450.00 |
| 5 | 38098.08 | 37901.38 | 38089.06 | 37985.40 | 37386.93 | 37253.76 | 37800.00 |
| 6 | 38071.13 | 38037.26 | 38100.16 | 38123.72 | 37414.91 | 37300.13 | 37400.00 |
| 7 | 38098.08 | 38001.35 | 38118.95 | 38018.36 | 37442.90 | 37293.41 | 37500.00 |
| 8 | 38123.29 | 38047.34 | 38145.84 | 38044.73 | 37470.91 | 37342.65 | 37700.00 |
| 9 | 38091.20 | 38034.37 | 38161.95 | 38057.55 | 37498.95 | 37371.35 | 36950.00 |
| 10 | 38174.92 | 38101.69 | 38191.55 | 38158.06 | 37527.01 | 37384.44 | 36750.00 |
| RMSE | 778.0506 | 718.0937 | 781.7677 | 776.3611 | 445.8601 | 458.239 | |
| MAPE | 0.01824928 | 0.01619924 | 0.01823542 | 0.01807621 | 0.008440755 | 0.0095819 | |

Bảng 3.12: So sánh hiệu suất dự báo với Giá thực

Nhận xét: Dự báo trực tiếp: **ARIMA(4,1,4)** tốt nhất với RMSE = 718.0937, MAPE = 0.0161; dự báo qua log return: **ARIMA(0,0,4)** tốt nhất với RMSE = 445.8601, MAPE = 0.00844. Như vậy, dự báo gián tiếp qua log return là tối ưu hơn. Chọn **ARIMA(0,0,4)** để dự báo giá CTG_t cho 10 phiên đầu năm 2025 (xem Bảng 3.12).

3.3 Phân tích chuỗi PVC - Nghiêm Gia Phương

3.3.1 Tổng quan về Tổng Công ty Hóa chất và Dịch vụ dầu khí - CTCP

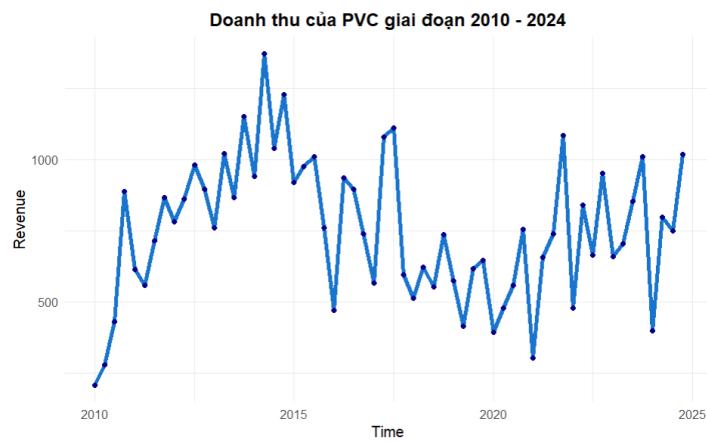
Tổng Công ty Hóa chất và Dịch vụ Dầu khí - Công ty Cổ phần (PVC), tiền thân là công ty Dung dịch khoan và Hoá phẩm Dầu khí được thành lập vào ngày 8/3/1990. Năm 2005, Công ty chuyển đổi mô hình kinh doanh hoạt động là Công ty Cổ phần. Ngày 15/11/2007, PVC chính thức giao dịch trên Sở Giao dịch Chứng khoán Hà Nội (HNX).

Là thành viên của Tổng công ty dầu khí Việt Nam, công ty chuyên về sản xuất và kinh doanh hoá chất, hoá phẩm; khai thác và chế biến khoáng sản phục vụ cho công nghiệp dầu khí và các ngành công nghiệp khác. Hiện nay, công ty là nhà cung cấp các dịch vụ hóa kỹ thuật hàng đầu cho các đơn vị trong và ngoài ngành dầu khí.

Dữ liệu được sử dụng trong bài nghiên cứu bao gồm 499 quan sát về giá cổ phiếu của mã PVC, được thu thập từ ngày 03/01/2023 đến ngày 31/12/2024. Từ các mức giá này, 498 giá trị log-return (lợi suất tính theo lãi suất liên tục) đã được tính toán nhằm phục vụ cho việc phân tích định lượng. Bên cạnh đó, dữ liệu theo quý bao gồm 60 quan sát về doanh thu, bắt đầu từ quý I năm 2010 đến quý IV năm 2024. Nguồn dữ liệu về doanh thu được thu thập từ báo cáo tài chính công bố trên trang FiinTrade, trong khi dữ liệu giá cổ phiếu được lấy từ trang Investing.

3.3.2 Dự báo doanh thu cho 4 quý năm 2025 của PVC

Trong giai đoạn 2010–2014, doanh thu của Tổng Công ty Hóa chất và Dịch vụ Dầu khí - CTCP ghi nhận mức tăng trưởng ấn tượng, phản ánh quá trình phát triển ổn định và bền vững của doanh nghiệp. Tuy nhiên, từ năm 2014 đến 2020, doanh thu có nhiều biến động và có xu hướng suy giảm, cho thấy những khó khăn mà công ty phải đối mặt trong giai đoạn này. Kể từ sau năm 2020, tình hình bắt đầu ổn định hơn và doanh thu dần hồi phục, đặc biệt rõ nét trong các năm 2023–2024. Điều này cho thấy doanh nghiệp đang từng bước thích nghi với điều kiện thị trường và dần lấy lại đà tăng trưởng. Đáng chú ý, doanh thu đạt đỉnh vào quý 2 năm 2014 với con số 1372.28 tỷ đồng; mặc dù có sự sụt giảm sau đó, mức doanh thu vẫn không thấp hơn thời điểm đầu kỳ, chứng tỏ công ty vẫn duy trì được quy mô hoạt động nhất định.



Hình 3.11: Đồ thị doanh thu của PVC giai đoạn 2010-2024

Các mô hình dự báo doanh thu bán hàng và cung cấp dịch vụ của Tổng Công ty Hóa chất và Dịch vụ dầu khí - CTCP với tập training set (2010 – 2023)

Ta sử dụng dữ liệu 56 quý, từ quý 1 năm 2010 đến quý 4 năm 2023 cho training set. Từ training set ta tạo ra các mô hình dự báo cho bốn quý năm 2024. So sánh dữ liệu được dự báo từ các mô hình với dữ liệu thực tế bằng RMSE và MAPE.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

Mô hình có RMSE và MAPE nhỏ nhất là mô hình dự báo chính xác nhất, sẽ được sử dụng để dự báo cho năm 2025.

Giải thích: cách tính ở mô hình (8) và mô hình (9):

- Mô hình (8): Holt - Winter có mùa vụ, dạng cộng:

$$DT_{T+k}^F = (934.19825 + 11.71475k) + S_q$$

với $S_1 = -296.51479, S_2 = -158.82354, S_3 = -105.98938, S_4 = 86.79820$

VD: Quý 3/2024: $DT_{56+3}^F = (934.19825 + 11.71475 \times 3) - 105.98938 = 863.3531$

- Mô hình (9): Holt - Winter có mùa vụ, dạng nhân:

$$DT_{T+k}^F = (1152.5983021 + 17.6979748k)S_q$$

với $S_1 = 0.5630960, S_2 = 0.6406033, S_3 = 0.7181630, S_4 = 0.8772290$

VD: Quý 2/2024: $DT_{56+2}^F = (1152.5983021 + 17.6979748 \times 2) \times 0.6406033 = 761.0330$

Dưới đây ta có bảng dự báo kết quả cho chuỗi doanh thu của Vietcombank trong năm 2024.

Lưu ý: dòng đầu là giá trị thực tế của năm 2024, từ dòng sau là tên mô hình, công thức và các giá trị dự báo.

| Quý 1 | Quý 2 | Quý 3 | Quý 4 | MAPE (%) | RMSE toàn mẫu | RMSE 2024 |
|---|----------|----------|-----------|-------------------|---------------|-----------|
| 397.872 | 796.804 | 749.63 | 1019.123 | (Giá trị thực tế) | | |
| (1) Mô hình lin - lin | | | | | | |
| $DT_t = 792.395 - 1.278t + u_t$ | | | | | | |
| 719.5593 | 718.2815 | 717.0037 | 715.7259 | 34.652 | 246.338 | 225.1461 |
| (2) Mô hình lin - log | | | | | | |
| $DT_t = 640.44 + 37.54\ln(t) + u_t$ | | | | | | |
| 792.2140 | 792.8669 | 793.5086 | 794.1395 | 34.466 | 244.9421 | 228.0703 |
| (3) Mô hình log - lin | | | | | | |
| $\ln(DT_t) = 6.572638 - 0.000270t + u_t$ | | | | | | |
| 704.3325 | 704.1424 | 703.9523 | 703.7623 | 33.187 | 251.248 | 225.8557 |
| (4) Mô hình log - log | | | | | | |
| $\ln(DT_t) = 6.2567 + 0.1002\ln(t) + u_t$ | | | | | | |
| 781.8271 | 783.1902 | 784.5324 | 785.8542 | 32.675 | 251.9493 | 225.4105 |
| (5) Hồi quy với biến giả mùa vụ | | | | | | |
| $DT_t = 584.28 + 187.21s_2 + 204.33s_3 + 295.24s_4 + u_t$ | | | | | | |
| 584.2826 | 771.4944 | 788.6102 | 879.5216 | 30.368 | 222.7039 | 118.7409 |
| (6) Xu thế tuyến tính + mùa vụ, dạng cộng | | | | | | |
| $DT_t = 630.669 - 1.718t + 188.930s_2 + 207.764s_3 + 300.393s_4 + u_t$ | | | | | | |
| 532.7421 | 719.9539 | 737.0696 | 827.9811 | 30.134 | 220.9742 | 123.2771 |
| (7) Xu thế tuyến tính + mùa vụ, dạng nhân | | | | | | |
| $DT_t = 799.186 - 6.658t + 5.549t \times s_2 + 6.033t \times s_3 + 8.530t \times s_4 + u_t$ | | | | | | |
| 419.6570 | 734.8252 | 762.3127 | 911.4601 | 30.977 | 224.8777 | 63.38019 |
| (8) Holt - Winter có mùa vụ, dạng cộng | | | | | | |
| $DT_{T+k}^F = (934.19825 + 11.71475k) + S_q$ | | | | | | |
| 649.3982 | 798.8042 | 863.3531 | 1067.8554 | 21.852 | 191.1194 | 140.1582 |
| (9) Holt - Winter có mùa vụ, dạng nhân | | | | | | |
| $DT_{T+k}^F = (1152.5983021 + 17.6979748k)S_q$ | | | | | | |
| 658.9892 | 761.0330 | 865.8835 | 1073.1934 | 22.479 | 195.1028 | 146.5437 |

Bảng 3.13: Các mô hình dự báo doanh thu của PVC

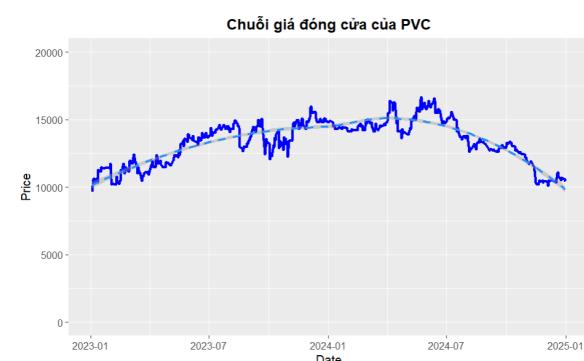
Dự báo doanh thu năm 2025

Vì mô hình xu thế tuyến tính + mùa vụ, dạng nhân có RMSE 2024 nhỏ nhất nên ta sẽ sử dụng mô hình này để dự báo doanh thu cho 4 quý năm 2025.

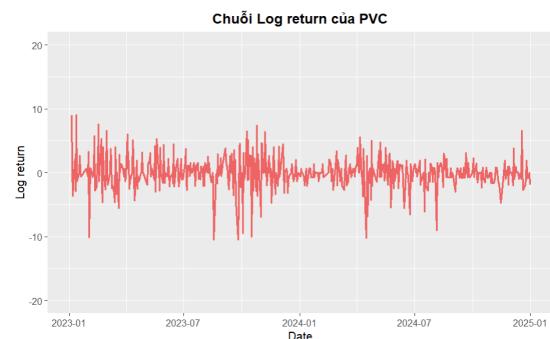
| | Quý 1 | Quý 2 | Quý 3 | Quý 4 |
|-------------------------|----------|----------|----------|----------|
| Kết quả dự báo năm 2025 | 390.9003 | 745.0881 | 759.6182 | 942.0749 |

Bảng 3.14: Kết quả dự báo doanh thu 4 quý năm 2025 của PVC

3.3.3 Dự báo giá cổ phiếu PVC 10 phiên đầu năm 2025



Hình 3.12: Đồ thị chuỗi giá đóng cửa PVC giai đoạn 2023 - 2024



Hình 3.13: Đồ thị chuỗi lợi suất PVC giai đoạn 2023 - 2024

Dựa trên Hình 3.12, chuỗi giá cổ phiếu PVC trong giai đoạn 2023–2024 có dấu hiệu của một chuỗi không dừng. Ở giai đoạn đầu, giá cổ phiếu tăng ổn định; tuy nhiên, sau khi đạt đỉnh vào khoảng giữa năm 2024, xu hướng bắt đầu đảo chiều và giảm rõ rệt đến hết kỳ quan sát.

Trong khi đó, Hình 3.13 thể hiện chuỗi lợi suất (log-return) của cổ phiếu PVC trong cùng giai đoạn. Quan sát cho thấy chuỗi này dao động xung quanh giá trị trung bình xấp xỉ bằng 0 và không có xu hướng rõ rệt theo thời gian. Điều này cho thấy chuỗi log-return có thể là một chuỗi dừng, phù hợp với lý thuyết tài chính rằng lợi suất thường ổn định hơn so với giá cổ phiếu. Tuy nhiên, chuỗi log-return vẫn nhận một số thời điểm biến động mạnh, đặc biệt là các đợt giảm sâu, cho thấy thị trường đôi lúc phản ứng tiêu cực. Dù vậy, biên độ dao động của chuỗi này có xu hướng thu hẹp dần về cuối giai đoạn.

* Dự báo 10 phiên cổ phiếu đầu tiên năm 2025

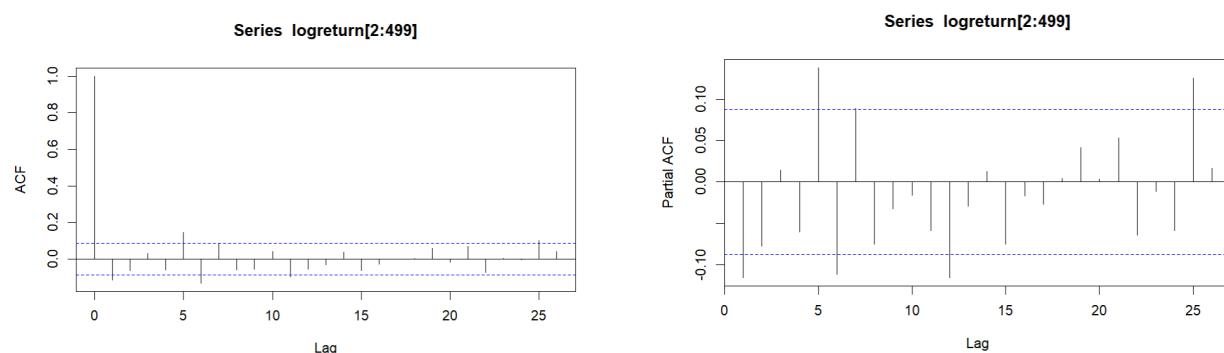
- **Bước 1: Kiểm định tính dừng:**

Bảng 3.15: Bảng kết quả kiểm định tính dừng ở mức ý nghĩa 5%

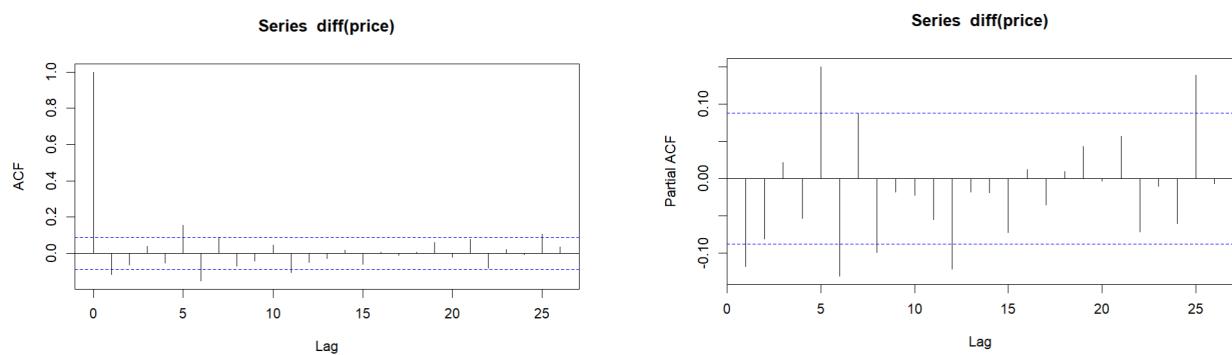
| Cấp giả thiết | $\text{Chuỗi } PVC_t$ | $\text{Chuỗi } LogReturnPVC_t$ | $\text{Chuỗi } \Delta PVC_t$ |
|--|---|--|---|
| Kiểm định ADF trend: H_0 : Chuỗi có nghiệm đơn vị (không dừng) H_1 : Chuỗi không có nghiệm đơn vị (dừng xu thế) | $ \tau_{stat} = 1.8169 < \tau_{0.05} = 3.42$ Chưa bác bỏ H_0 , chuỗi PVC_t không dừng. $p-value = 0.2114 > 0.05$ Hệ số xu thế tt không có ý nghĩa thống kê. | $ \tau_{stat} = 18.2422 > \tau_{0.05} = 3.42$ Bác bỏ H_0 , chuỗi $LogReturnPVC_t$ dừng. $p-value = 0.0729 > 0.05$ Hệ số xu thế tt không có ý nghĩa thống kê. | $ \tau_{stat} = 18.234 > \tau_{0.05} = 3.42$ Bác bỏ H_0 , chuỗi ΔPVC_t dừng. $p-value = 0.0962 > 0.05$ Hệ số xu thế tt không có ý nghĩa thống kê. |
| Kiểm định ADF drift: H_0 : Chuỗi có nghiệm đơn vị (không dừng) H_1 : Chuỗi không có nghiệm đơn vị (dừng vây quanh trung bình dài hạn) | $ \tau_{stat} = 2.0392 < \tau_{0.05} = 2.87$ Chưa bác bỏ H_0 , chuỗi PVC_t không dừng. $p-value = 0.0435 < 0.05$ Hệ số chặn có ý nghĩa thống kê. | $ \tau_{stat} = 18.1128 > \tau_{0.05} = 2.87$ Bác bỏ H_0 , chuỗi $LogReturnPVC_t$ dừng. $p-value = 0.9813 > 0.05$ Hệ số chặn không có ý nghĩa thống kê. | $ \tau_{stat} = 18.1252 > \tau_{0.05} = 2.87$ Bác bỏ H_0 , chuỗi ΔPVC_t dừng. $p-value = 0.9853 > 0.05$ Hệ số chặn không có ý nghĩa thống kê. |
| Kiểm định ADF none: H_0 : Chuỗi có nghiệm đơn vị (không dừng) H_1 : Chuỗi không có nghiệm đơn vị (dừng vây quanh 0) | | $ \tau_{stat} = 18.1317 > \tau_{0.05} = 1.95$ Bác bỏ H_0 , chuỗi $LogReturnPVC_t$ dừng. | $ \tau_{stat} = 18.1439 > \tau_{0.05} = 1.95$ Bác bỏ H_0 , chuỗi ΔPVC_t dừng. |

Vậy $PVC_t \sim I(1)$ và $LogReturnPVC_t \sim I(0)$

- **Bước 2: Chọn mô hình ARIMA phù hợp để ước lượng, kiểm tra, đánh giá và dự báo**

Hình 3.14: Biểu đồ ACF và PACF của chuỗi $LogReturnPVC_t$

Đề xuất mô hình ARIMA(1,0,5) và ARIMA(5,0,1) cho chuỗi $LogReturnPVC_t$

Hình 3.15: Biểu đồ ACF và PACF của chuỗi ΔPVC_t

Đề xuất mô hình ARIMA(1,1,5) và ARIMA(5,1,1) cho chuỗi PVC_t

Bảng 3.16: So sánh các mô hình ARIMA

Bảng A: Chuỗi LogReturnPVC_t

| Thông số | ARIMA(1,0,5) | ARIMA(5,0,1) |
|-------------------------|--------------|----------------|
| Hệ số | | |
| ar1 | -0.6044** | -0.7550** |
| ar2 | — | -0.1632** |
| ar3 | — | -0.0349 |
| ar4 | — | -0.0366 |
| ar5 | — | 0.1159** |
| ma1 | 0.5231** | 0.6701** |
| ma2 | -0.1516** | — |
| ma3 | -0.0110 | — |
| ma4 | -0.0144 | — |
| ma5 | 0.1297** | — |
| Chỉ tiêu mô hình | | |
| AIC | 2322.76 | 2320.41 |

Bảng B: Chuỗi PVC_t

| Thông số | ARIMA(1,1,5) | ARIMA(5,1,1) |
|-------------------------|--------------|----------------|
| Hệ số | | |
| ar1 | -0.6261** | -0.7880** |
| ar2 | — | -0.1709** |
| ar3 | — | -0.0290 |
| ar4 | — | -0.0218 |
| ar5 | — | 0.1303** |
| ma1 | 0.5545** | 0.7102** |
| ma2 | -0.1566 ** | — |
| ma3 | 0.0000 | — |
| ma4 | -0.0036 | — |
| ma5 | 0.1396** | — |
| Chỉ tiêu mô hình | | |
| AIC | 7179.07 | 7176.62 |

Ghi chú: ** $p < 0.05$; * $p < 0.1$. Giá trị AIC càng nhỏ càng tốt.

Nhận xét: Trong Bảng A, mô hình ARIMA(5,0,1) có giá trị AIC thấp hơn so với mô hình ARIMA(1,0,5), vì vậy mô hình ARIMA(5,0,1) được lựa chọn để dự báo chuỗi LogReturnPVC_t. Tương tự, trong Bảng B, mô hình ARIMA(5,1,1) cho kết quả AIC thấp hơn so với mô hình ARIMA(1,1,5), do đó mô hình ARIMA(5,1,1) được sử dụng để dự báo chuỗi PVC_t.

- Bước 3: Kiểm định về mô hình: Kết quả phân tích được trình bày dưới phần phụ lục cho thấy tất cả các mô hình ARIMA đã đề xuất đều có nghiệm nghịch đảo nằm trong đường tròn đơn vị, không có tự tương quan giữa các bậc, hay phần dư là nhiễu trắng.

- Bước 4: Dự báo cho 10 phiên đầu năm 2025:

Bảng 3.17: Dự báo giá cổ phiếu PVC 10 phiên đầu năm 2025

| Phiên | Log-return dự báo (%) | Giá dự báo | Giá thực | Phiên | Giá dự báo | Giá thực |
|-----------------|-----------------------|------------|----------|-----------------|------------|----------|
| 1 | 0.3949 | 10441.15 | 10500 | 1 | 10438.50 | 10500 |
| 2 | 0.0470 | 10446.06 | 10500 | 2 | 10444.52 | 10500 |
| 3 | -0.1423 | 10431.21 | 10300 | 3 | 10425.96 | 10300 |
| 4 | 0.1556 | 10447.45 | 10100 | 4 | 10442.79 | 10100 |
| 5 | -0.3311 | 10412.92 | 10200 | 5 | 10405.63 | 10200 |
| 6 | 0.2737 | 10441.45 | 10100 | 6 | 10437.46 | 10100 |
| 7 | -0.1474 | 10426.07 | 10100 | 7 | 10419.43 | 10100 |
| 8 | 0.0560 | 10431.91 | 10200 | 8 | 10426.49 | 10200 |
| 9 | 0.0024 | 10432.16 | 10200 | 9 | 10426.09 | 10200 |
| 10 | -0.0542 | 10426.51 | 10500 | 10 | 10420.18 | 10500 |
| RMSE | 1.0544 | 229.2686 | | RMSE | 225.2126 | |
| MAPE (%) | Inf | 1.97 | | MAPE (%) | 1.945 | |

(a) Dự báo giá gián tiếp qua mô hình ARIMA(5,0,1) của chuỗi LogReturnPVC_t (b) Dự báo giá trực tiếp qua mô hình ARIMA(5,1,1)

Nhận xét: So với phương pháp dự báo giá gián tiếp từ chuỗi LogReturnPVC_t thông qua mô hình ARIMA(5,0,1), phương pháp dự báo giá trực tiếp với mô hình ARIMA(5,1,1) cho kết quả gần với giá thực tế hơn.

3.3.4 Dự báo rủi ro cho 10 phiên đầu năm 2025 bằng mô hình GARCH

Từ kết quả dự báo giá cổ phiếu PVC 10 phiên đầu năm 2025, ta đã kết luận được rằng mô hình ARIMA(5,1,1) là mô hình dự báo tốt nhất cho chuỗi PVC_t. Do vậy ta sẽ sử dụng mô hình trung bình ở đây là ARMA(5,1) cho chuỗi dừng LogReturnPVC_t để dự báo rủi ro cho 10 phiên này. Ngoài ra, ta đề xuất mô hình Garch(1,1) để dự đoán rủi ro cho 10 phiên đầu năm 2025, mô hình Garch(1,1) có dạng:

$$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + \alpha_1 u_{t-1}^2 + \nu_t$$

Ước lượng mô hình được kết quả:

| mu | ar1 | ar2 | ar3 | ar4 | ar5 | ma1 | omega |
|-------------|-------------|-------------|-------------|-------------|-------------|------------|------------|
| -0.06073097 | -1.01168222 | -0.18879876 | -0.08041550 | -0.04941130 | -0.02197967 | 0.96646594 | 0.17070152 |
| alpha1 | beta1 | | | | | | |
| 0.07233003 | 0.89813929 | | | | | | |

Hình 3.16: Kết quả mô hình Garch

| Phiên | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| σ | 208.5859 | 209.9275 | 211.2214 | 212.4695 | 213.6738 | 214.8360 | 215.9580 | 217.0413 | 218.0874 | 219.0979 |

Bảng 3.18: Dự báo rủi ro của giá cổ phiếu PVC 10 phiên đầu năm 2025

3.4 Phân tích chuỗi PVS - Nguyễn Thị Trà

3.4.1 Giới thiệu

Tổng Công ty Cổ phần Dịch vụ Kỹ thuật Dầu khí Việt Nam (PTSC) là một trong những đơn vị chủ lực của Tập đoàn Dầu khí Quốc gia Việt Nam (PVN). Được thành lập từ tháng 2 năm 1993, PTSC hoạt động chủ yếu trong lĩnh vực cung cấp dịch vụ kỹ thuật cho ngành dầu khí và công nghiệp, chuyên về các loại dịch vụ như: tổng thầu thiết kế - cung ứng - thi công - lắp đặt (EPCI); tổng thầu EPC công trình biển và công nghiệp; vận hành kho nổ xử lý và xuất dầu thô (FSO/FPSO); cùng các dịch vụ tàu chuyên dụng phục vụ khảo sát, vận chuyển và hỗ trợ ngoài khơi.

PTSC chính thức cổ phần hóa vào năm 2007 và hiện đang niêm yết trên Sở Giao dịch Chứng khoán Hà Nội (HNX) với mã chứng khoán PVS. Về cơ cấu cổ đông, ba cổ đông lớn nhất hiện nay gồm: Tập đoàn Dầu khí Việt Nam (PVN), CTBC Vietnam Equity Fund, và Công ty TNHH MTV Quản lý Quỹ Ngân hàng TMCP Công Thương Việt Nam (VietinBank Capital). Với tiềm lực mạnh mẽ và chiến lược phát triển rõ ràng, PVS là một trong những mã cổ phiếu dầu khí tiêu biểu thu hút sự quan tâm của giới đầu tư trên thị trường chứng khoán Việt Nam.

Dữ liệu được sử dụng trong bài nghiên cứu bao gồm 499 quan sát về giá cổ phiếu của mã PVS, được thu thập từ ngày 03/01/2023 đến ngày 31/12/2024, tương ứng 498 giá trị log-return (lợi suất tính theo lãi suất liên tục). Bên cạnh đó, dữ liệu theo quý bao gồm 60 quan sát về doanh thu, bắt đầu từ quý I năm 2010 đến quý IV năm 2024. Nguồn dữ liệu về doanh thu được thu thập từ báo cáo tài chính công bố trên FiinTrade, dữ liệu giá cổ phiếu được lấy từ trang CafeF.

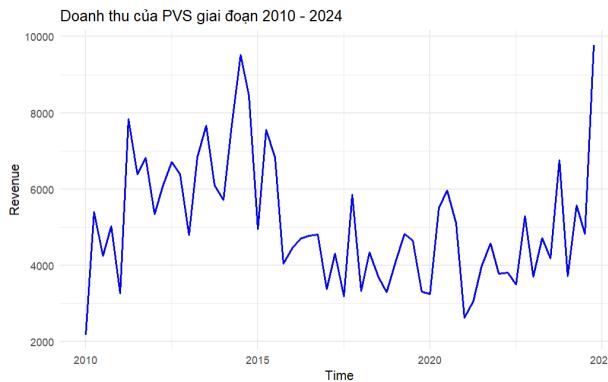
3.4.2 Dự báo chuỗi doanh thu năm 2025

Doanh thu của PVS trong giai đoạn 2010-2024 cho thấy một xu hướng biến động rõ rệt theo thời gian. Giai đoạn 2010-2015 là thời kỳ doanh thu tăng trưởng mạnh, đạt đỉnh vào quý 3 năm 2014 với hơn 9.500 tỷ đồng, phản ánh sự phát triển nhanh chóng nhưng cũng nhiều biến động. Từ năm 2016 đến 2019, doanh thu giảm sút và duy trì ở mức thấp hơn, dao động chủ yếu trong khoảng 4.000–6.000, cho thấy sự chững lại trong hoạt động kinh doanh, do ảnh hưởng từ các yếu tố vĩ mô như giá dầu thế giới không ổn định, các sự kiện chính trị thế giới. Giai đoạn 2020–2023 tiếp tục ghi nhận sự dao động mạnh trong ngắn hạn, phản ánh những ảnh hưởng tiêu cực từ đại dịch COVID-19 và các biến động kinh tế toàn cầu. Tuy nhiên, cuối năm 2024, doanh thu của PVS tăng vọt trở lại, đạt mức cao nhất trong toàn bộ chuỗi thời gian, cho thấy tín hiệu phục hồi rõ rệt và tiềm năng tăng trưởng trở lại trong tương lai gần.

Các mô hình dự báo doanh thu của PVS với tập training set (2010 – 2023)

Ta sử dụng dữ liệu 56 quý, từ quý 1 năm 2010 đến quý 4 năm 2023 cho training set. Từ training set, ta tạo ra các mô hình dự báo cho 4 quý năm 2024. Đối với mô hình sử dụng biến giả mùa vụ, chọn quý 1 làm cơ sở. So sánh dữ liệu được dự báo từ các mô hình với dữ liệu thực tế bằng RMSE và MAPE và sử dụng mô hình có RMSE và MAPE nhỏ nhất để dự báo cho năm 2025.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} ; \quad \text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$



Hình 3.17: Doanh thu của PVS giai đoạn 2010-2024

Dưới đây ta có bảng dự báo kết quả cho chuỗi doanh thu của PVS trong năm 2024.

| Quý 1 | Quý 2 | Quý 3 | Quý 4 | MAPE trong mẫu | RMSE trong mẫu | RMSE 2024 |
|--|----------|----------|----------|-------------------|-------------------|--------------|
| 3710.321 | 5571.760 | 4821.429 | 9776.991 | (Giá trị thực tế) | | |
| (1) Mô hình lin - lin | | | | | | |
| $DT_t = 6103.56 - 37.10t + u_t$ | | | | | | |
| 3989.114 | 3952.018 | 3914.923 | 3877.827 | 0.2535988 | 1455.182 | 3095.286 |
| (2) Mô hình lin - log | | | | | | |
| $DT_t = 6116.9 + -347.8ln(t) + u_t$ | | | | | | |
| 4710.569 | 4704.519 | 4698.573 | 4692.727 | 0.2693596 | 1543.227 | 2627.614 |
| (3) Mô hình log - lin | | | | | | |
| $ln(DT_t) = 8.670377 - 0.006739t + u_t$ | | | | | | |
| 3968.920 | 3942.263 | 3915.785 | 3889.485 | 0.2401093 | 1474.816 | 3090.512 |
| (4) Mô hình log - log | | | | | | |
| $ln(DT_t) = 8.64269 - 0.05341ln(t) + u_t$ | | | | | | |
| 4567.661 | 4563.420 | 4559.256 | 4555.165 | 0.2530201 | 1566.279 | 2696.664 |
| (5) Hồi quy với dạng bậc hai của trend | | | | | | |
| $DT_T = 5799.6468 - 5.6566t - 0.5516t^2 + u_t$ | | | | | | |
| 3685.205 | 3616.119 | 3545.930 | 3474.638 | 0.2558426 | 1449.469 | 3360.495 |
| (6) Hồi quy với biến giả mùa vụ | | | | | | |
| $DT_t = 3914.9 + 1562.2s_2 + 1463.2s_3 + 1500.5s_4 + u_t$ | | | | | | |
| 3914.862 | 5477.109 | 5378.052 | 5415.315 | 0.2526173 | 1431.453 | 2201.411 |
| (7) Xu thế tuyến tính + mùa vụ, dạng cộng | | | | | | |
| $DT_t = 4978.40 - 39.39t + 1601.64s_2 + 1541.97s_3 + 1618.62s_4 + u_t$ | | | | | | |
| 2733.153 | 4295.400 | 4196.342 | 4233.606 | 0.2241215 | 1282.826 | 2902.749 |
| (8) Xu thế tuyến tính + mùa vụ, dạng nhân | | | | | | |
| $DT_t = 6134.50 - 64.04t + 30.81t \times s_2 + 30.57t \times s_3 + 39.94t \times s_4 + u_t$ | | | | | | |
| 2484.420 | 4207.435 | 4160.048 | 4688.760 | 0.2331992 | 1372.045 | 2724.506 |
| (9) Xu thế không tuyến tính + mùa vụ, dạng cộng | | | | | | |
| $ln(DT_t) = 8.434027 - 0.007223t + 0.340363s_2 + 0.314129s_3 + 0.346116s_4 + u_t$ | | | | | | |
| 3048.167 | 4253.235 | 4113.288 | 4216.421 | 0.2147455 | 1287.034 | 2898.204 |
| (10) Xu thế không tuyến tính + mùa vụ, dạng nhân | | | | | | |
| $ln(DT_t) = 8.676903 - 0.012493t + 0.006749t \times s_2 + 0.006376t \times s_3 + 0.008534t \times s_4 + u_t$ | | | | | | |
| 2877.812 | 4203.664 | 4088.713 | 4625.427 | 0.222745 | 1386.987 | 2722.142 |
| (11) Holt - Winter có mùa vụ, dạng cộng | | | | | | |
| $DT_{T+k}^F = (5547.0896 + 129.8870k) + S_q$ | | | | | | |
| với $S_1 = -1292.5699, S_2 = -430.3554, S_3 = -438.6385, S_4 = 575.6250$ | | | | | | |
| 4384.407 | 5376.508 | 5498.112 | 6642.263 | 0.2289383 | 1330.723 | 1641.412 |
| (12) Holt - Winter có mùa vụ, dạng nhân | | | | | | |
| $DT_{T+k}^F = (5908.7863119 + 134.0998790k)S_q$ | | | | | | |
| với $S_1 = 0.6991583, S_2 = 0.848001, S_3 = 0.8457521, S_4 = 1.0799589$ | | | | | | |
| 4224.934 | 5238.095 | 5337.614 | 6960.536 | 0.2211784 | 1386.374 | 1464.157 |

Bảng 3.19: Bảng kết quả dự báo cho chuỗi doanh thu PVS trong năm 2024

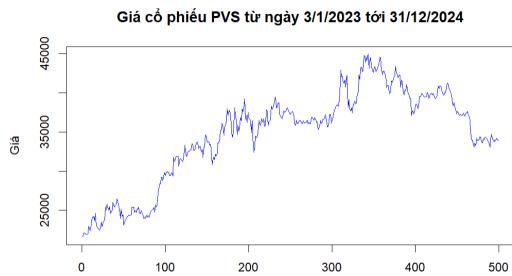
Dự báo doanh thu năm 2025

Từ kết quả của bảng 3.19, ta có mô hình (7) và (9) có RMSE và MAPE trên tập mẫu nhỏ nhất. Tuy nhiên, mô hình Holt - Winter có mùa vụ dạng nhân có MAPE và RMSE trên tập validation nhỏ nhất nên ta sẽ dùng mô hình này để dự báo cho năm 2025.

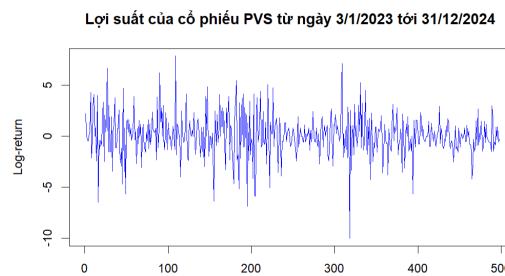
| | Quý 1 | Quý 2 | Quý 3 | Quý 4 |
|-------------------------|----------|----------|----------|----------|
| Kết quả dự báo năm 2025 | 4903.955 | 5896.056 | 6017.660 | 7161.811 |

Bảng 3.20: Kết quả dự báo cho chuỗi doanh thu PVS trong năm 2025

3.4.3 Dự báo giá cổ phiếu 10 phiên đầu năm 2025



Hình 3.18: Chuỗi giá cổ phiếu PVS



Hình 3.19: Chuỗi log-return của cổ phiếu PVS

Giá cổ phiếu PVS nhìn chung không dừng và tồn tại xu hướng tăng - giảm rõ rệt theo thời gian. Năm 2023, giá cổ phiếu tăng đều đặn, phản ánh sự kỳ vọng tích cực từ phía nhà đầu tư. Giá cổ phiếu sau đó dao động mạnh, đạt đỉnh gần 45000 đồng vào khoảng giữa năm 2024 và giảm nhẹ vào giai đoạn cuối năm.

Ngược lại, chuỗi lợi suất log-return có thể được xem là chuỗi dừng về mặt trực quan. Lợi suất log dao động quanh mức 0, với một số giai đoạn có biên độ biến động mạnh, vượt ngưỡng $\pm 5\%$.

Để có kết luận chắc chắn về tính dừng của 2 chuỗi này, cần sử dụng các kiểm định.

Xây dựng mô hình dự báo 10 phiên cổ phiếu đầu tiên năm 2025

- **Bước 1: Kiểm định tính dừng:** Tiến hành kiểm tra tính dừng của chuỗi giá và chuỗi log-return bằng kiểm định ADF lần lượt với:

Kiểm định "trend": $\Delta Y_t = \mu + \delta Y_{t-1} + \beta t + u_t$

$$\begin{cases} H_0 : \delta = 0 : \text{Chuỗi có nghiệm đơn vị (Chuỗi không dừng)} \\ H_1 : \delta < 0 : \text{Chuỗi không có nghiệm đơn vị (Chuỗi dừng xu thế)} \end{cases}$$

Kiểm định "drift": $\Delta Y_t = \mu + \delta Y_{t-1} + u_t$

$$\begin{cases} H_0 : \delta = 0 : \text{Chuỗi có nghiệm đơn vị (Chuỗi không dừng)} \\ H_1 : \delta < 0 : \text{Chuỗi không có nghiệm đơn vị (Chuỗi dừng quanh } \mu^*) \end{cases}$$

Kiểm định "none": $\Delta Y_t = \delta Y_{t-1} + u_t$

$$\begin{cases} H_0 : \delta = 0 : \text{Chuỗi có nghiệm đơn vị (Chuỗi không dừng)} \\ H_1 : \delta < 0 : \text{Chuỗi không có nghiệm đơn vị (Chuỗi dừng quanh 0)} \end{cases}$$

Ta được kết quả sau:

Bảng 3.21: Kết quả kiểm định tính dừng của chuỗi giá PVS

| | P-value | $ \tau_{stat} $ | $ \tau_\alpha $ | Với mức ý nghĩa 5% | Kết luận |
|-------|---------|-----------------|-----------------|--------------------|---|
| trend | 0.9659 | 1.3918 | 3.42 | Chưa bác bỏ H_0 | Hệ số của t không có ý nghĩa thống kê, PVS_t không dừng |
| drift | 0.0226 | 2.1755 | 2.87 | Chưa bác bỏ H_0 | Hệ số chặn có ý nghĩa thống kê, PVS_t không dừng |
| none | | 0.4719 | 1.95 | Chưa bác bỏ H_0 | PVS_t không dừng |

Do chuỗi giá PVS_t không là chuỗi dừng với mức ý nghĩa 5%, ta tiếp tục thực hiện kiểm định ADF đối với chuỗi sai phân giá PVS.

Bảng 3.22: Kết quả kiểm định tính dừng của chuỗi sai phân giá PVS

| | P-value | $ \tau_{stat} $ | $ \tau_\alpha $ | Với mức ý nghĩa 5% | Kết luận |
|-------|---------|-----------------|-----------------|--------------------|---|
| trend | 0.0824 | 17.4555 | 3.42 | Bác bỏ H_0 | Hệ số của t không có ý nghĩa thống kê, ΔPVS_t dừng xu thế |
| drift | 0.381 | 17.3329 | 2.87 | Bác bỏ H_0 | Hệ số chặn không có ý nghĩa thống kê |
| none | | 17.3148 | 1.95 | Bác bỏ H_0 | ΔPVS_t dừng quanh 0 |

Với mức ý nghĩa 5%, chuỗi $\Delta PVS_t \sim I(0)$, hay $PVS_t \sim I(1)$.

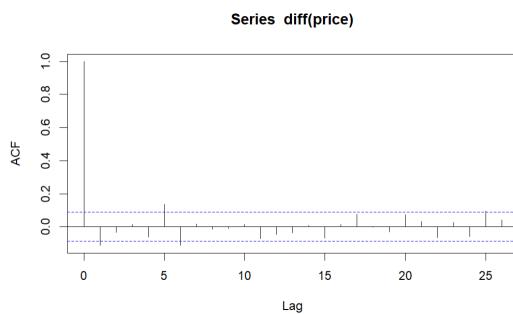
Bảng 3.23: Kết quả kiểm định tính dừng của chuỗi log-return giá PVS

| | P-value | $ \tau_{stat} $ | $ \tau_\alpha $ | Với mức ý nghĩa 5% | Kết luận |
|-------|---------|-----------------|-----------------|--------------------|---|
| trend | 0.0491 | 17.1943 | 3.42 | Bắc bỏ H_0 | Hệ số của t có ý nghĩa thống kê, $\log PVS_t$ dừng xu thế |
| drift | 0.279 | 17.0309 | 2.87 | Bắc bỏ H_0 | Hệ số chặn không có ý nghĩa thống kê |
| none | | 16.9934 | 1.95 | Bắc bỏ H_0 | $\log PVS_t$ dừng quanh 0 |

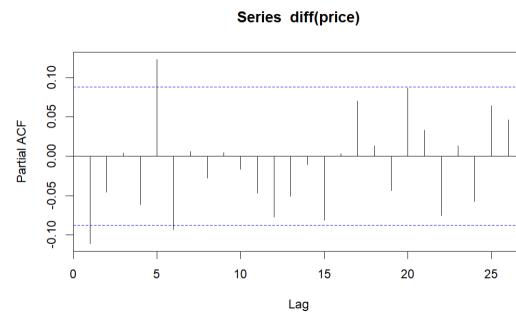
Như vậy, chuỗi log-return giá PVS là chuỗi dừng quanh 0 với mức ý nghĩa 5%.

- **Bước 2: Chọn mô hình ARIMA phù hợp để ước lượng, kiểm tra, đánh giá và dự báo Chuỗi giá PVS**

Trước tiên, ta cần xác định bậc $ARIMA(p, d, q)$ cho chuỗi giá PVS. Do $PVS_t \sim I(1)$ nên hệ số d của mô hình ARIMA phù hợp với chuỗi giá PVS_t là 1. Ta cần xác định bậc $AR(p)$ và $MA(q)$.



Hình 3.20: ACF của ΔPVS_t



Hình 3.21: PACF của ΔPVS_t

Nhìn vào đồ thị ACF của chuỗi ΔPVS_t , ta thấy các hệ số tương quan khác không ở các độ trễ 1, 5 và 6. Còn đối với biểu đồ PACF, ta có các hệ số tương quan riêng phần khác nhau ở các độ trễ 1, 5 và 6. Tạm thời ta xác định độ trễ cao nhất của p là 6 và độ trễ cao nhất của q là 6. Ta đề xuất một số mô hình ARIMA dưới đây. (Xem thêm kết quả chi tiết ở phần phụ lục).

Mô hình phù hợp là mô hình có tối thiểu hoá các chỉ tiêu AIC, BIC, RMSE, các biến có ý nghĩa thống kê và càng ít độ trễ càng tốt.

Với mức ý nghĩa 5%, ta so sánh $|t_{qs}| = |\frac{coef}{se}|$ với $\alpha_{0.05} = 1.96$. Nếu $|t_{qs}| > 1.96$ thì hệ số đó có ý nghĩa thống kê. Thử lần lượt các mô hình, xem xét đổi sang mô hình bậc thấp hơn khi nhiều hệ số không có ý nghĩa thống kê.

Bảng 3.24: Kết quả mô hình ARIMA cho chuỗi giá đóng cửa PVS

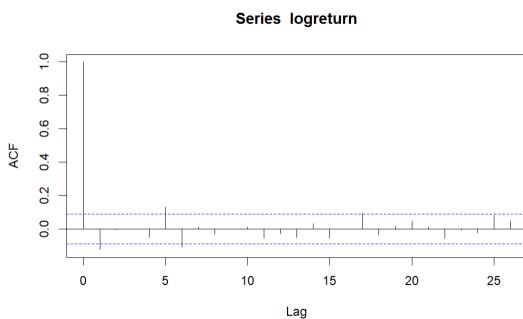
| Chỉ tiêu đánh giá | AIC | BIC | RMSE | Hệ số có ý nghĩa |
|-------------------|----------------|----------------|-----------------|--|
| Mô hình | | | | Mức ý nghĩa 5% |
| ARIMA(6,1,1) | 7939.21 | 7972.9 | 688.7875 | none |
| ARIMA(6,1,0) | 7937.28 | 7966.75 | 688.8327 | ar1, ar5, ar6 |
| ARIMA(5,1,6) | 7945.79 | 7996.32 | 687.7916 | none |
| ARIMA(5,1,5) | 7941.27 | 7987.59 | 685.9422 | ar1, ar4, ar5 , ma4, ma5 |
| ARIMA(5,1,1) | 7937.6 | 7967.07 | 689.0554 | ar1, ma1 |
| ARIMA(1,1,6) | 7939.83 | 7973.51 | 689.2215 | ma5 |
| ARIMA(1,1,5) | 7939.3 | 7968.78 | 690.2583 | ar1 , ma1, ma5 |
| ARIMA(1,1,1) | 7943.04 | 7943.09 | 698.5202 | none |

*Lưu ý: none là không có hệ số có ý nghĩa thống kê với mức ý nghĩa 5 %

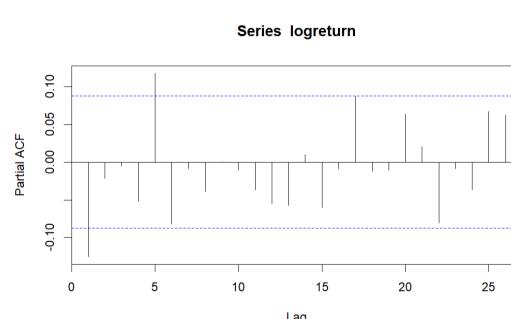
Từ kết quả bảng 3.24, ta lựa chọn mô hình ARIMA(6,1,0), ARIMA(1,1,5), ARIMA(5,1,5) để tiếp tục kiểm tra, đánh giá và dự báo.

Chuỗi log-return giá PVS

Ta thấy $\log PVS_t \sim I(0)$ nên cần xác định bậc ARMA(p, q) của chuỗi $\log PVS_t$.



Hình 3.22: ACF của $\log PVS_t$



Hình 3.23: PACF của $\log PVS_t$

Tương tự với chuỗi giá, qua đồ thị ACF và PACF của chuỗi $\log PVS_t$, ta đề xuất các mô hình ARIMA sau: (Xem thêm kết quả chi tiết ở phần phụ lục).

Bảng 3.25: Kết quả mô hình ARIMA cho chuỗi log-return giá PVS

| Chỉ tiêu đánh giá | AIC | BIC | RMSE | Hệ số có ý nghĩa |
|-------------------|----------------|----------------|-----------------|-----------------------------------|
| Mô hình | | | | Mức ý nghĩa 5% |
| ARIMA(5,0,6) | 2120.24 | 2170.76 | 1.980971 | ar3, ma1, ma2, ma4, ma6 |
| ARIMA(5,0,5) | 2119.82 | 2166.14 | 1.984966 | ar3, ar5 , ma3, ma5 |
| ARIMA(5,0,1) | 2118.81 | 2127.23 | 2.022537 | ar1, ar5 , ma1 |
| ARIMA(1,0,6) | 2119.34 | 2153.02 | 1.999133 | ma5 |
| ARIMA(1,0,5) | 2118.58 | 2148.05 | 2.001661 | ar1 , ma1, ma5 |
| ARIMA(1,0,1) | 2120.62 | 2133.25 | 2.022148 | none |

*Lưu ý: none là không có hệ số có ý nghĩa thống kê với mức ý nghĩa 5 %

Ta lựa chọn mô hình ARIMA(1,0,5), ARIMA(5,0,1), ARIMA(5,0,5) cho bước tiếp theo.

- **Bước 3: Kiểm định về mô hình:** Tiến hành kiểm định tính nhiễu trắng của phần dư, xem tính dừng qua nghiệm nghịch đảo. (Xem thêm kết quả chi tiết ở phần phụ lục).

Từ kết quả, ta thấy các mô hình ARIMA cho chuỗi giá đóng cửa đã đề xuất đều có nghiệm nghịch đảo nằm trong đường tròn đơn vị, không có tự tương quan giữa các bậc, hay phần dư là nhiễu trắng với mức ý nghĩa 5%.

Với mô hình ARIMA cho chuỗi log-return, mô hình ARIMA(5,0,5) cho nghiệm nghịch đảo nằm trên đường tròn đơn vị. Ở mô hình ARIMA(1,0,5), ARIMA(5,0,1), ACF phần dư có lồi nhẹ ở bậc 16, vượt ra ngoài ngưỡng tin cậy trong đồ thị ACF. Tuy nhiên, AR bậc cao dễ gây overfitting và sau khi kiểm định cả hai mô hình đều có $p - value$ rất cao (> 0.9), cho thấy phần dư không có tự tương quan, hay phần dư là nhiễu trắng nên ta vẫn chấp nhận mô hình.

- **Bước 4: Dự báo cho 10 phiên đầu năm 2025:**

Ta có được kết quả dự báo chuỗi giá đóng cửa PVS cho 10 phiên đầu năm 2025 từ mô hình ARIMA cho chuỗi giá, bao gồm ARIMA(6,1,0), ARIMA(1,1,5), ARIMA(5,1,5).

Ví dụ cụ thể dạng của mô hình: ARIMA(6,1,0): $\Delta PVS_t = -0.0953\Delta PVS_{t-1} - 0.0501\Delta PVS_{t-2} + 0.0051\Delta PVS_{t-3} - 0.0494\Delta PVS_{t-4} + 0.1142\Delta PVS_{t-5} - 0.0907\Delta PVS_{t-6}$

Với mô hình ARIMA cho chuỗi log-return, ta suy ra chuỗi giá theo công thức sau để so sánh với chuỗi giá thực:

$$PVS_t = PVS_{t-1} \times e^{\frac{\log PVS_t}{100}}$$

Đối với phiên thứ 500, ta sử dụng giá trị cuối cùng trong chuỗi giá đóng cửa của dữ liệu thực từ 2023 - 2024 là PVS_{499} để tính cho PVS_{500}^F . (Xem thêm kết quả chuyển đổi cụ thể ở phần phụ lục).

Dưới đây là kết quả dự báo bằng mô hình ARIMA cho chuỗi giá và log-return sau khi cùng đưa về giá đóng cửa, sử dụng tiêu chí RMSE nhỏ nhất để đánh giá mô hình

Bảng 3.26: So sánh kết quả chuỗi giá được tính từ mô hình cho chuỗi giá và log-return

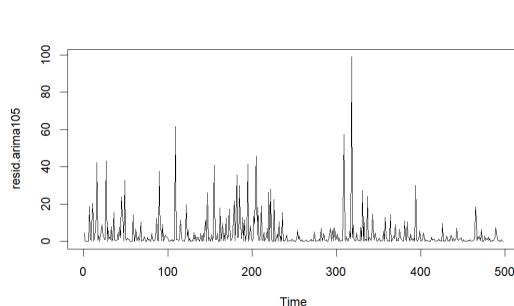
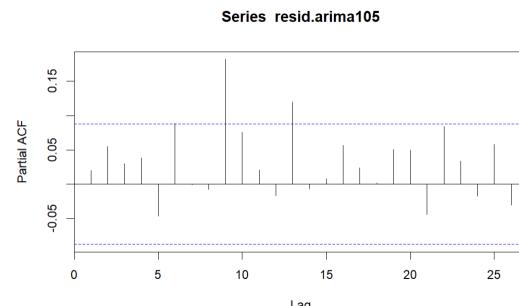
| Phiên | Chuỗi giá PVS | | | Chuỗi log-return | | Giá thực |
|-------|---------------|--------------|--------------|------------------|--------------|----------|
| | ARIMA(6,1,0) | ARIMA(1,1,5) | ARIMA(5,1,5) | ARIMA(1,0,5) | ARIMA(5,0,1) | |
| 500 | 33987.51 | 33963.07 | 33987.24 | 33950.63 | 33972.23 | 33800 |
| 501 | 33929.70 | 33910.46 | 33864.15 | 33906.15 | 33919.34 | 34000 |
| 502 | 33983.51 | 33980.52 | 33946.83 | 33971.67 | 33984.00 | 33400 |
| 503 | 33936.62 | 33929.43 | 33937.72 | 33921.67 | 33932.54 | 32100 |
| 504 | 33940.50 | 33947.78 | 33951.70 | 33939.53 | 33945.77 | 32100 |
| 505 | 33964.68 | 33939.03 | 33931.10 | 33930.27 | 33950.28 | 32200 |
| 506 | 33944.74 | 33943.20 | 33910.72 | 33935.07 | 33939.63 | 32000 |
| 507 | 33959.15 | 33941.22 | 33970.51 | 33932.58 | 33953.79 | 32600 |
| 508 | 33948.47 | 33942.16 | 33947.02 | 33933.87 | 33940.53 | 32100 |
| 509 | 33952.17 | 33941.71 | 33938.41 | 33933.20 | 33948.56 | 31700 |
| RMSE | 1561.208 | 1554.05 | 1552.623 | 1546.77 | 1557.059 | |

Nhận xét: Phương pháp dự báo giá gián tiếp từ chuỗi log $PV S_t$ thông qua mô hình ARIMA(1,0,5) cho kết quả gần với giá thực tế hơn.

3.4.4 Mô hình ARCH/GARCH và các dự báo

Từ kết quả dự báo giá cổ phiếu PVS 10 phiên đầu năm 2025, ta đã kết luận được rằng mô hình ARIMA(1,0,5) là mô hình dự báo tốt nhất cho chuỗi log $PV S_t$. Do vậy ta sẽ sử dụng mô hình với trung bình ở đây là ARMA(1,5) cho chuỗi dừng log $PV S_t$ để dự báo rủi ro cho 10 phiên này. Mô hình tổng quát là:

$$\log PV S_t = ARMA(1, 5) + u_t$$

Hình 3.24: Đồ thị biểu diễn u_t^2 Hình 3.25: PACF của u_t^2

Đồ thị bình phương phần dư của chuỗi log $PV S_t$ cho thấy phần dư biến động không đồng đều.

Đồ thị PCAF của u_t^2 cho thấy u_t^2 có thể có tự tương quan ở bậc cao. Kiểm định tự động ARCH(6), ARCH(9), ARCH(14) cho thấy có ARCH ở bậc tương ứng với mức ý nghĩa 5%. Do vậy, ta đề xuất sử dụng mô hình GARCH(1,1) do mô hình ARCH(p) tương đương GARCH(1,1) khi $p \rightarrow \infty$.

Ước lượng mô hình GARCH(1,1) ta có kết quả như sau:

$$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + \alpha_1 u_{t-1}^2 + v_t \quad \text{với } \alpha_0 = 0.099440^*, \beta_1 = 0.908313^{***}, \alpha_1 = 0.068401^{***}$$

Với mức ý nghĩa 5%, các hệ số đều có ý nghĩa thống kê; đồng thời, kiểm định Jarque-Bera cho thấy phần dư không tuân theo phân phối chuẩn, và kiểm định Ljung-Box cho thấy phần dư không có tự tương quan, hay phần dư là nhiễu trắng. Vậy nên, ta chấp nhận và tiến hành sử dụng mô hình GARCH(1,1) để dự báo độ rủi ro cho 10 phiên đầu năm 2025.

Bảng 3.27: Dự báo rủi ro cho 10 phiên đầu năm 2025

| Phiên | 500 | 501 | 502 | 503 | 504 | 505 | 506 | 507 | 508 | 509 |
|----------|--------|---------|--------|---------|--------|---------|--------|---------|--------|--------|
| σ | 0.1753 | -0.1087 | 0.1679 | -0.1358 | 0.0698 | -0.0261 | 0.0268 | -0.0024 | 0.0137 | 0.0048 |

Chương 4

BÀI NHÓM

4.1 Xây dựng danh mục tối ưu

4.1.1 Phân tích tổng quan

Trung bình

| | VCB | CTG | PVC | PVS |
|--------------|-----------|-----------|-----------|-----------|
| Giá cổ phiếu | 86,891.82 | 30,177.42 | 13,365.64 | 34,667.78 |
| Lợi suất | 0.0533 | 0.0826 | 0.0152 | 0.0907 |

Bảng 4.1: Bảng giá trị trung bình của giá cổ phiếu và lợi suất các chuỗi

Bảng 4.1 cho thấy giá trị trung bình của giá cổ phiếu và lợi suất tương ứng của bốn mã cổ phiếu: VCB, CTG, PVC và PVS. Trong đó, VCB có mức giá cổ phiếu trung bình cao nhất (khoảng 87 nghìn đồng), tiếp theo là CTG (khoảng 30 nghìn đồng), PVS (khoảng 35 nghìn đồng) và thấp nhất là PVC (khoảng 13,4 nghìn đồng). Điều này phản ánh sự chênh lệch rõ rệt về quy mô và giá trị thị trường giữa các doanh nghiệp được niêm yết.

Về mặt lợi suất, cổ phiếu PVS có mức lợi suất trung bình cao nhất (0.0907), tiếp theo là CTG (0.0826), VCB (0.0533) và thấp nhất là PVC (0.0152). Mặc dù giá cổ phiếu PVC thấp nhất, nhưng nó cũng mang lại lợi suất thấp nhất, trong khi PVS có thể là mã cổ phiếu hấp dẫn hơn xét theo lợi suất kỳ vọng trung bình.

Từ bảng số liệu này, nhà đầu tư có thể cân nhắc giữa việc đầu tư vào cổ phiếu có giá cao hơn nhưng lợi suất trung bình (như VCB) hay các cổ phiếu có mức giá vừa phải nhưng lợi suất hấp dẫn hơn (như PVS).

Hệ số tương quan

| | VCB | CTG | PVC | PVS |
|-----|----------|----------|----------|-----|
| VCB | 1 | | | |
| CTG | 0.770222 | 1 | | |
| PVC | 0.396424 | 0.095332 | 1 | |
| PVS | 0.763834 | 0.632337 | 0.691446 | 1 |

Hình 4.1: Hệ số tương quan giữa các chuỗi giá

| | rVCB | rCTG | rPVC | rPVS |
|------|----------|----------|----------|------|
| rVCB | 1 | | | |
| rCTG | 0.320591 | 1 | | |
| rPVC | 0.195542 | 0.388939 | 1 | |
| rPVS | 0.22752 | 0.436216 | 0.804259 | 1 |

Hình 4.2: Hệ số tương quan giữa các chuỗi lợi suất

Dựa trên bảng hệ số tương quan giữa các chuỗi giá cổ phiếu (VCB, CTG, PVC và PVS), có thể thấy phần lớn các cặp biến có mối tương quan dương khá cao, phản ánh xu hướng đồng biến giữa các chuỗi giá trong giai đoạn nghiên cứu. Đặc biệt, do cùng nhóm ngành, các cặp VCB – CTG và

VCB – PVS có hệ số tương quan lần lượt là 0.770 và 0.768, cho thấy mối liên hệ chặt chẽ giữa biến động giá của các mã cổ phiếu này. Ngược lại hai cổ phiếu CTG và PVC, lần lượt thuộc nhóm ngành ngân hàng và dầu khí, lại có hệ số tương quan rất thấp (0.095), cho thấy gần như không có mối liên hệ trong biến động giá giữa chúng.

So với chuỗi giá, hệ số tương quan giữa các chuỗi lợi suất thể hiện xu hướng thấp hơn đáng kể, phản ánh tính biến động riêng biệt và yếu tố ngẫu nhiên lớn hơn trong lợi suất ngắn hạn của từng cổ phiếu. Trong đó, chỉ có cặp $r_{PVC} - r_{PVS}$ đạt hệ số tương quan cao (0.804), cho thấy lợi suất của hai cổ phiếu này có xu hướng biến động cùng chiều mạnh mẽ, có thể do ngành dầu khí chịu tác động chung. Các cặp còn lại có hệ số dao động từ khoảng 0.20 đến 0.43, phản ánh mức độ tương quan yếu đến trung bình, từ đó cho thấy việc phân tích rủi ro và đa dạng hóa danh mục đầu tư thông qua lợi suất sẽ mang lại kết quả khác biệt so với phân tích dựa trên giá cổ phiếu.

Phương sai

| | <i>VCB</i> | <i>CTG</i> | <i>PVC</i> | <i>PVS</i> |
|------------|------------|------------|------------|------------|
| <i>VCB</i> | 35122147 | | | |
| <i>CTG</i> | 18521899 | 16464874 | | |
| <i>PVC</i> | 3815418 | 628215.2 | 2637440 | |
| <i>PVS</i> | 26632346 | 15095504 | 6606471 | 34613066 |

Hình 4.3: Hiệp phương sai giữa các chuỗi giá

| | <i>r_{VCB}</i> | <i>r_{CTG}</i> | <i>r_{PVC}</i> | <i>r_{PVS}</i> |
|------------------------|------------------------|------------------------|------------------------|------------------------|
| <i>r_{VCB}</i> | 1.442733 | | | |
| <i>r_{CTG}</i> | 0.625107 | 2.635234 | | |
| <i>r_{PVC}</i> | 0.594841 | 1.599042 | 6.414119 | |
| <i>r_{PVS}</i> | 0.556445 | 1.441856 | 4.147387 | 4.145914 |

Hình 4.4: Hiệp phương sai giữa các chuỗi lợi suất

Ma trận phương sai – hiệp phương sai giữa các chuỗi giá cho thấy mức độ biến động và mối quan hệ tuyến tính giữa các cặp cổ phiếu. Có thể thấy giá cổ phiếu VCB và PVS có phương sai sai lầm nhất, lần lượt là 35.1 triệu và 34.6 triệu, cho thấy mức độ dao động giá của hai mã này cao hơn các mã còn lại trong giai đoạn nghiên cứu. Ngoài ra, các cặp của PVC đều có hiệp phương sai thấp nhất trong bảng, cho thấy cặp cổ phiếu này có xu hướng biến động cùng chiều yếu hơn so với các cặp còn lại.

Trong khi đó, ma trận phương sai – hiệp phương sai của các chuỗi lợi suất thể hiện mức độ biến động thấp hơn nhiều so với chuỗi giá, phản ánh đặc trưng thường thấy của lợi suất được tính theo sự thay đổi tỉ lệ. Phương sai cao nhất thuộc về r_{PVC} và r_{PVS} (lần lượt là 6.41 và 4.14), cho thấy hai mã này có lợi suất dao động lớn hơn các mã khác, hàm ý rủi ro đầu tư cao hơn. Các giá trị hiệp phương sai dương như giữa $r_{PVC} - r_{PVS}$ (4.14) và $r_{CTG} - r_{PVC}$ (1.59) phản ánh sự đồng biến lợi suất giữa các mã, trong khi những giá trị thấp hơn như $r_{VCB} - r_{CTG}$ (0.62) cho thấy mức độ liên hệ yếu hơn. Điều này rất quan trọng trong việc xây dựng danh mục đầu tư nhằm phân tán rủi ro thông qua lựa chọn các cổ phiếu có mối liên hệ lợi suất thấp.

4.1.2 Tối thiểu hóa phương sai

Đây là một trong những khái niệm quan trọng trong lý thuyết tối ưu hóa danh mục đầu tư của Markowitz. Danh mục có phương sai nhỏ đặc biệt hữu ích với đa số nhà đầu tư duy lý bởi họ là những người sợ rủi ro. Họ là những người muốn giảm thiểu rủi ro và bảo vệ vốn đầu tư của mình trong các điều kiện thị trường không chắc chắn.

Bài toán này được thể hiện như sau: Xây dựng một danh mục gồm log - return của 4 cổ phiếu VCB, CTG, PVC, PVS sao cho phương sai danh mục là nhỏ nhất, với các điều kiện như tỉ trọng của từng tài sản không nhỏ hơn 0 và không lớn hơn 1; tổng các tỉ trọng cần tìm phải bằng 1; giá trị kỳ vọng của danh mục không nhỏ hơn trung bình lợi suất của thị trường.

Mục tiêu của bài toán: $V(r_P) \rightarrow \min$

Các điều kiện:

- $0 \leq w_i \leq 1$ với $i = 1, 2, 3, 4$

- $\sum_{i=1}^4 w_i = 1$
- $E(r_P) \geq E(r_{market})$

Trong đó:

$$V(r_P) = \sigma_P^2 = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \text{Cov}(r_i, r_j) = W' V W$$

W là véc-tơ tỉ trọng của các tài sản hay ma trận tỉ trọng của các tài sản gồm n hàng ứng với n tài sản và 1 cột, V là ma trận hiệp phương sai của các tài sản, W' là ma trận chuyển vị của W .

$$E(r_P) = \sum_{i=1}^N w_i E(r_i) = W' E(r)$$

W' là ma trận chuyển vị của ma trận tỉ trọng các cổ phiếu, $E(r)$ là ma trận giá trị lợi suất kì vọng của các cổ phiếu.

$$E(r_{market}) = 0.038976$$

Trung bình lợi suất thị trường được tính bằng cách lấy giá trị trung bình lợi suất VNIndex trong giai đoạn 2023-2024.

Giải bài toán trên, ta thu được các kết quả như sau:

| Tỉ trọng theo cổ phiếu | VCB | CTG | PVC | PVS |
|----------------------------|-----------------|----------|-----|----------|
| | 0.665450 | 0.220389 | 0 | 0.114161 |
| Phương sai nhỏ nhất | 1.161358 | | | |
| Kì vọng danh mục | 0.064015 | | | |

Bảng 4.2: Kết quả phương án tối thiểu hóa phương sai danh mục

Trong trường hợp tối thiểu hóa phương sai, kết quả cho thấy danh mục đầu tư được phân bổ chủ yếu vào cổ phiếu VCB với tỉ trọng 66.55%, tiếp theo là CTG chiếm 22.04% và PVS chiếm 11.42%. Đáng chú ý, cổ phiếu PVC không được lựa chọn trong danh mục do cổ phiếu này có lợi suất trung bình thấp và phương sai lớn, do đó không mang lại lợi ích đáng kể. Phương sai tối thiểu đạt được là 1.161358, cho thấy rõ ràng tổng thể của danh mục đã được kiểm soát hiệu quả nhờ vào sự phân tán đầu tư vào các tài sản có mức độ tương quan hợp lý. Đồng thời, lợi suất kỳ vọng của danh mục là 6.4%, cao hơn so với lợi suất trung bình của thị trường là 3.9%. Điều này phản ánh rằng việc xây dựng danh mục theo nguyên tắc tối thiểu hóa rủi ro không những giúp giảm thiểu biến động mà vẫn đảm bảo hiệu quả đầu tư tốt, phù hợp với những nhà đầu tư có mức độ chấp nhận rủi ro thấp nhưng vẫn mong muốn thu được mức sinh lời hấp dẫn.

4.1.3 Tối đa hóa lợi nhuận

Trong quá trình ra quyết định đầu tư, bên cạnh mục tiêu kiểm soát rủi ro, nhiều nhà đầu tư lại ưu tiên tối đa hóa lợi nhuận thu được từ danh mục của mình. Khác với bài toán tối thiểu hóa phương sai, phương pháp tiếp cận này thiên về tối đa hóa phần thu về và chấp nhận khả năng biến động lớn hơn. Việc giải bài toán tối đa hóa lợi nhuận giúp nhà đầu tư xây dựng được danh mục có hiệu suất cao nhất, phù hợp với khẩu vị rủi ro cao hoặc mục tiêu sinh lời mạnh trong ngắn hạn.

Bài toán tối đa hóa lợi nhuận nhằm mục tiêu tìm ra tổ hợp phân bổ tỉ trọng giữa các tài sản sao cho lợi suất kỳ vọng của danh mục là cao nhất, đồng thời vẫn thỏa mãn các điều kiện ràng buộc nhất định về giới hạn tỉ trọng (tương tự bài toán tối thiểu hóa phương sai) và mức độ rủi ro không vượt quá rủi ro thị trường.

Mục tiêu của bài toán: $E(r_P) \rightarrow \max$

Các điều kiện:

- $0 \leq w_i \leq 1$ với $i = 1, 2, 3, 4$
- $\sum_{i=1}^4 w_i = 1$
- $V(r_P) \leq V(r_{market})$

Trong đó các công thức tính $E(r_P)$ và $V(r_P)$ đã được trình bày trong bài toán tối thiểu hóa rủi ro ở trên.

$$V(r_{market}) = 0.917860$$

Dữ liệu phương sai thị trường được tính từ VNIndex trong giai đoạn 2023-2024. Tuy nhiên, do phương sai của thị trường chỉ là 0.917860, nhỏ hơn phương sai của tất cả các cổ phiếu, bài toán này vô nghiệm.

Do đó, ta đổi điều kiện, thay phương sai thị trường bằng trung bình nhân của các phương sai mỗi cổ phiếu."

$$V(r_p) = \sqrt[4]{1.44273 \times 2.63523 \times 6.41412 \times 4.14591} = 3.170959$$

Khi đó, ta có bài toán mới:

Mục tiêu của bài toán: $E(r_P) \rightarrow \max$

Các điều kiện:

- $0 \leq w_i \leq 1$ với $i = 1, 2, 3, 4$
- $\sum_{i=1}^4 w_i = 1$
- $V(r_P) \leq 3.170959$

Giải bài toán trên, ta thu được các kết quả như sau:

| Tỉ trọng theo cổ phiếu | VCB | CTG | PVC | PVS |
|----------------------------------|-----------------|---------|-----|---------|
| | 0 | 0.21296 | 0 | 0.78704 |
| Lợi suất kỳ vọng lớn nhất | 0.088981 | | | |
| Phương sai danh mục | 3.170958 | | | |

Bảng 4.3: Kết quả phương án tối đa hóa lợi suất kỳ vọng danh mục

Trong quá trình tối ưu hóa nhằm tối đa hóa lợi suất kỳ vọng của danh mục, mô hình đã phân bổ tỉ trọng hoàn toàn cho hai cổ phiếu là CTG (21.296%) và đặc biệt là PVS (78.704%), trong khi VCB và PVC không được lựa chọn. Điều này phản ánh rằng hai cổ phiếu CTG và PVS mang lại mức lợi suất kỳ vọng cao hơn, từ đó có đóng góp lớn hơn vào mục tiêu tối đa hóa lợi suất danh mục. Lợi suất kỳ vọng của danh mục đạt được là 0.088981, vượt mức trung bình của các cổ phiếu riêng lẻ, cho thấy hiệu quả của việc phân bổ có chọn lọc. Tuy nhiên, đi kèm với lợi suất cao là phương sai danh mục ở mức 3.170958, phản ánh rủi ro của danh mục cũng tăng lên đáng kể so với phương án tối thiểu hóa phương sai.

4.2 Dự báo chuỗi lợi suất danh mục

Trong phân tích danh mục đầu tư, việc dự báo chuỗi lợi suất trong tương lai đóng vai trò quan trọng nhằm hỗ trợ nhà đầu tư đưa ra các quyết định phân bổ tài sản hợp lý. Phần dưới đây sẽ trình bày hai cách dự báo chuỗi lợi suất của danh mục:

- 1, dự báo chuỗi lợi suất của từng cổ phiếu, sau đó nhân chuỗi lợi suất dự báo vừa tìm được với tỉ trọng tìm được từ phương án tối thiểu hóa phương sai.
- 2, dự báo cho tích của chuỗi lợi suất gốc của từng cổ phiếu với tỉ trọng tìm được từ phương án tối thiểu hóa phương sai.

4.2.1 Từ chuỗi lợi suất gốc của danh mục

Ở phần này, ta sẽ dự báo lợi suất danh mục dựa trên chuỗi lợi suất của từng tài sản cấu thành danh mục đó.

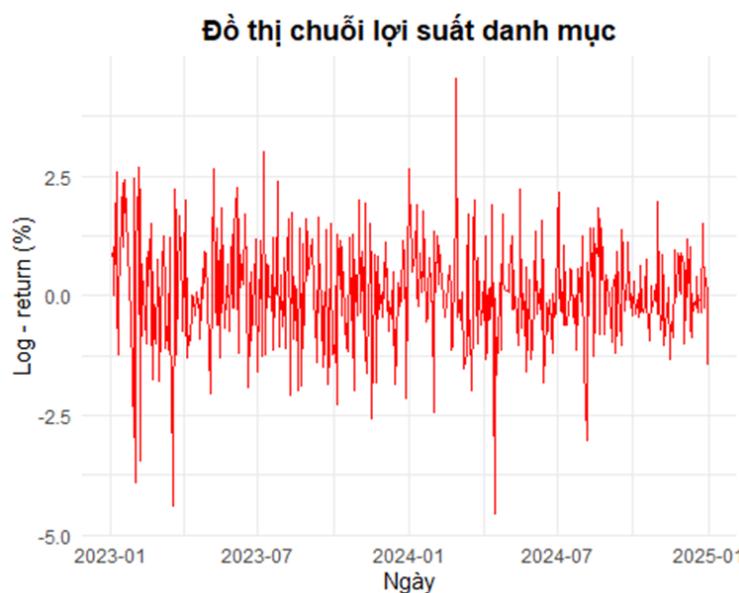
Bước quan trọng trong quá trình dự báo này là xây dựng chuỗi lợi suất danh mục dựa trên chuỗi lợi suất gốc đã được thu thập. Cách thực hiện là nhân ma trận chuỗi lợi suất với ma trận tỉ trọng.

$$\begin{pmatrix} r_{1,VCB} & r_{1,CTG} & r_{1,PVC} & r_{1,PVS} \\ r_{2,VCB} & r_{2,CTG} & r_{2,PVC} & r_{2,PVS} \\ \vdots & \vdots & \vdots & \vdots \\ r_{498,VCB} & r_{498,CTG} & r_{498,PVC} & r_{498,PVS} \end{pmatrix} \cdot \begin{pmatrix} w_{VCB} \\ w_{CTG} \\ w_{PVC} \\ w_{PVS} \end{pmatrix} = \begin{pmatrix} r_{1,p} \\ r_{2,p} \\ \vdots \\ r_{498,p} \end{pmatrix}$$

Ma trận chuỗi lợi suất gồm 498 dòng tương ứng với 498 quan sát lợi suất, 4 cột tương ứng với 4 mã cổ phiếu khác nhau. Ma trận tỉ trọng có kích thước 4×1 , chính là tỉ trọng của từng cổ phiếu tìm được từ phương án tối thiểu hóa phương sai (mục 4.1.2). Ma trận kết quả là chuỗi lợi suất danh mục cần tìm, gồm 498 dòng tương ứng với 498 quan sát.

Ta có thể thay giá trị thực tế vào và thu được kết quả chuỗi lợi suất danh mục:

$$\begin{pmatrix} 0.24134 & 1.77011 & 8.96117 & 2.20927 \\ 1.43875 & 0 & 0.71178 & 0.43425 \\ \vdots & \vdots & \vdots & \vdots \\ -1.19892 & -2.86852 & -1.90482 & -0.29455 \end{pmatrix} \cdot \begin{pmatrix} 0.66544 \\ 0.22039 \\ 0 \\ 0.11416 \end{pmatrix} = \begin{pmatrix} 0.80293 \\ 1.00699 \\ \vdots \\ -1.46364 \end{pmatrix}$$



Hình 4.5: Đồ thị chuỗi lợi suất danh mục giai đoạn 2023-2024

Biểu đồ chuỗi lợi suất danh mục cho thấy lợi suất dao động mạnh quanh mức 0% và có sự biến động khá lớn trong suốt giai đoạn quan sát. Đặc biệt, các cú sốc lợi suất xuất hiện nhiều hơn vào đầu các năm, phản ánh tính bất ổn của thị trường trong những khoảng thời gian này. Mức độ biến động có xu hướng giảm dần về cuối kỳ. Chuỗi không thể hiện xu hướng rõ ràng, phù hợp với đặc điểm của chuỗi lợi suất tài chính.

Dự báo lợi suất 10 phiên đầu tiên năm 2025

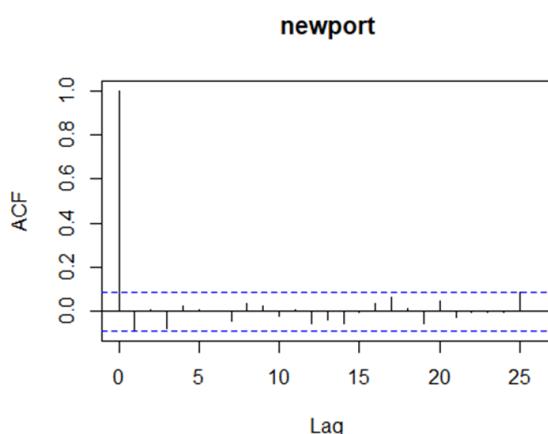
- Bước 1: Kiểm định tính dừng:

| Cặp giả thiết | Kết quả ở mức ý nghĩa 5% |
|---|---|
| Kiểm định ADF có xu thế: $\left\{ \begin{array}{l} H_0 : \text{Chuỗi có nghiệm đơn vị (Chuỗi không dừng)} \\ H_1 : \text{Chuỗi không có nghiệm đơn vị (Chuỗi dừng xu thế)} \end{array} \right.$ | $ \tau_{qs} = 16.4629 > \tau_{0.05} = 3.42$ Bắc bỏ H_0 , chuỗi log - return VCB dừng. $p - value = 0.294 > 0.05$ Hệ số xu thế tt không có ý nghĩa thống kê. |
| Kiểm định ADF chỉ có hệ số chặn: $\left\{ \begin{array}{l} H_0 : \text{Chuỗi có nghiệm đơn vị (Chuỗi không dừng)} \\ H_1 : \text{Chuỗi không có nghiệm đơn vị (Chuỗi dừng vây quanh TB dài hạn)} \end{array} \right.$ | $ \tau_{qs} = 16.4276 > \tau_{0.05} = 2.87$ Bắc bỏ H_0 , chuỗi log - return VCB dừng. $p - value = 0.170 > 0.05$ Hệ số chặn không có ý nghĩa thống kê. |
| Kiểm định ADF không có hệ số chặn: $\left\{ \begin{array}{l} H_0 : \text{Chuỗi có nghiệm đơn vị (Chuỗi không dừng)} \\ H_1 : \text{Chuỗi không có nghiệm đơn vị (Chuỗi dừng vây quanh 0)} \end{array} \right.$ | $ \tau_{qs} = 16.3558 > \tau_{0.05} = 1.952$ Bắc bỏ H_0 , chuỗi log - return VCB dừng. $\tau_{qs} = -16.3558 < 0$ Không cần thêm yếu tố hệ số chặn hoặc xu thế. |

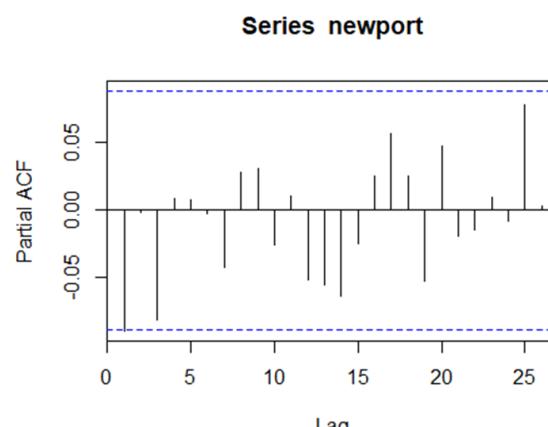
Bảng 4.4: Bảng kết quả kiểm định tính dừng của chuỗi lợi suất danh mục

Kết luận: Từ kết quả trên, ta nhận thấy chuỗi lợi suất danh mục là chuỗi dừng vây quanh 0.

- Bước 2: Chọn mô hình ARIMA phù hợp để ước lượng, kiểm tra, đánh giá và dự báo:



Hình 4.6: Biểu đồ ACF chuỗi lợi suất danh mục



Hình 4.7: Biểu đồ PACF chuỗi lợi suất danh mục

Từ 2 đồ thị trên, có thể thấy tương quan giữa biến lợi suất danh mục với trễ tại bậc 1 và tương quan với shock của 1 kì trước đó. Ta đề xuất mô hình ARIMA(1,0,1) và ARIMA(1,0,0).

| Thông số | ARIMA(1,0,1) | ARIMA(1,0,0) |
|------------------|--------------|----------------|
| Hệ số | | |
| ar1 | 0.4654 | -0.0894* |
| ma1 | -0.5503 | - |
| Chỉ tiêu mô hình | | |
| AIC | 1491.23 | 1489.78 |

Bảng 4.5: Bảng so sánh các mô hình ARIMA cho chuỗi lợi suất danh mục

Ghi chú: * $p < 0.05$, hệ số có ý nghĩa thống kê.

Mô hình ARIMA(1,0,0) có hệ số AIC thấp nhất, đồng thời hệ số ar1 cũng có ý nghĩa thống kê nên ta chọn mô hình này để dự báo chuỗi lợi suất của danh mục.

- Bước 3: Kiểm định về mô hình: Kết quả phân tích được trình bày dưới phần phụ lục cho thấy rằng tất cả các mô hình ARIMA sử dụng đều có nghiệm nghịch đảo nằm trong đường tròn đơn vị không có tự tương quan giữa các bậc, hay phần dư là nhiễu trắng.

- Bước 4: Dự báo cho 10 phiên đầu năm 2025: Từ các mô hình trên, ta có được kết quả dự báo chuỗi lợi suất danh mục 10 phiên đầu năm 2025. Ta suy ra chuỗi giá theo công thức sau để so sánh với chuỗi giá thực:

$$P_t = P_{t-1} \times e^{\frac{\log P_t}{100}}$$

| Phiên | Log - return dự báo (%) | Giá quy đổi từ log-return (nghìn VND) | Giá thực (nghìn VND) |
|-------------|-------------------------|---------------------------------------|----------------------|
| 1 | 0.2008 | 73036.05 | 72845.44 |
| 2 | 0.0519 | 73074.00 | 72914.94 |
| 3 | 0.0652 | 73121.68 | 73052.32 |
| 4 | 0.0640 | 73168.53 | 73073.58 |
| 5 | 0.0642 | 73215.49 | 73106.23 |
| 6 | 0.0641 | 73262.47 | 73134.47 |
| 7 | 0.0641 | 73309.48 | 73176.82 |
| 8 | 0.0641 | 73356.52 | 73217.09 |
| 9 | 0.0641 | 73403.59 | 73257.39 |
| 10 | 0.0641 | 73450.69 | 73296.07 |
| RMSE | | 136.419 | |
| MAPE | | 0.1811% | |

Bảng 4.6: Log-return dự báo, giá quy đổi và giá thực tế của danh mục trong 10 phiên đầu năm 2025

4.2.2 Từ kết quả chuỗi dự báo cá nhân

Với các trọng số tính toán ở bước trước, ta kết hợp với các giá trị log-return dự báo từ mô hình ARIMA trong bài cá nhân để được chuỗi lợi suất của portfolio:

| Phiên | r_{VCB} | r_{CTG} | r_{PVC} | r_{PVS} | Portfolio |
|-------|-----------|-----------|-----------|-----------|------------|
| 1 | 0.1224 | 0.0946 | 0.3695 | 0.1492 | 0.11934322 |
| 2 | 0.1111 | 0.0857 | 0.0577 | -0.1311 | 0.07782490 |
| 3 | 0.2005 | 0.0957 | -0.1779 | 0.1931 | 0.17653155 |
| 4 | 0.0323 | 0.0884 | 0.1613 | -0.1473 | 0.02415085 |
| 5 | 0.0394 | 0.0801 | -0.3565 | 0.0526 | 0.04986874 |
| 6 | 0.0371 | 0.0801 | 0.3054 | -0.0273 | 0.03926244 |
| 7 | 0.0577 | 0.0801 | -0.1729 | 0.0142 | 0.05764876 |
| 8 | 0.0556 | 0.0801 | 0.0677 | -0.0073 | 0.05383365 |
| 9 | 0.0549 | 0.0801 | -0.0038 | 0.0038 | 0.05462356 |
| 10 | 0.0525 | 0.0801 | -0.0567 | -0.0020 | 0.05238267 |

Bảng 4.7: Chuỗi log-return dự báo của các cổ phiếu và portfolio theo phiên

4.2.3 So sánh kết quả dự báo lợi suất chuỗi portfolio từ 2 cách

Ta có bảng giá trị thực chuỗi lợi suất của 4 mã cổ phiếu trong 10 phiên đầu của năm 2025.

| Phiên | r_{VCB} | r_{CTG} | r_{PVC} | r_{PVS} |
|-------|-----------|-----------|-----------|-----------|
| 1 | 0.1224 | 0.0946 | 0.3695 | 0.1492 |
| 2 | 0.1111 | 0.0857 | 0.0577 | -0.1311 |
| 3 | 0.2005 | 0.0957 | -0.1779 | 0.1931 |
| 4 | 0.0323 | 0.0884 | 0.1613 | -0.1473 |
| 5 | 0.0394 | 0.0801 | -0.3565 | 0.0526 |
| 6 | 0.0371 | 0.0801 | 0.3054 | -0.0273 |
| 7 | 0.0577 | 0.0801 | -0.1729 | 0.0142 |
| 8 | 0.0556 | 0.0801 | 0.0677 | -0.0073 |
| 9 | 0.0549 | 0.0801 | -0.0038 | 0.0038 |
| 10 | 0.0525 | 0.0801 | -0.0567 | -0.0020 |

Bảng 4.8: Dữ liệu chuỗi lợi suất thực

So sánh kết quả dự báo chuỗi lợi suất portfolio theo 2 cách:

| Phiên | Cách 1 | Cách 2 | Giá trị thực |
|-------|------------|------------------|--------------|
| 1 | 0.20067282 | 0.11934322 | 0.6493 |
| 2 | 0.05194476 | 0.07782490 | -0.3871 |
| 3 | 0.06523556 | 0.17653155 | 0.4446 |
| 4 | 0.06404785 | 0.02415085 | -0.7368 |
| 5 | 0.06415399 | 0.04986874 | 0.2771 |
| 6 | 0.06414450 | 0.03926244 | -0.5600 |
| 7 | 0.06414535 | 0.05764876 | -0.3753 |
| 8 | 0.06414527 | 0.05383365 | 0.6923 |
| 9 | 0.06414528 | 0.05462356 | -0.7643 |
| 10 | 0.06414528 | 0.05238267 | -0.4808 |
| RMSE | 0.5643792 | 0.5575605 | |

Bảng 4.9: So sánh hai cách theo phiên

Từ kết quả trên, ta có thể thấy cách tính chuỗi lợi suất portfolio theo cách 2 cho kết quả gần với dữ liệu thực hơn. Do đó, đây được coi là cách dự báo hiệu quả hơn.

4.3 Phân tích đồng tích hợp

Đồng tích hợp là hiện tượng hai hay nhiều chuỗi thời gian có tồn tại một số nguyên nhân gốc làm chúng có xu hướng biến động tương đồng nhau về dài hạn ngay cả khi nhìn riêng lẻ thì chúng chỉ là bước ngẫu nhiên.

Do các chuỗi mà nhóm sử dụng đều là chuỗi không dừng, trong khi đó lại dừng ở sai phân bậc nhất, nên cần kiểm định khả năng xảy ra các vector đồng tích hợp giữa các chuỗi thời gian này.Kỹ thuật kiểm định đồng tích hợp được áp dụng là phương pháp của Johansen và Juselius được đề xuất vào năm 1990.Đây là kỹ thuật kiểm định đồng tích hợp phổ biến nhất trong việc áp dụng nguyên tắc cực đại hợp lý, nhằm xác định sự tồn tại của các vector đồng tích hợp giữa các chuỗi thời gian không dừng. Phương pháp này có ưu điểm là có thể giúp xác định số lượng các vector đồng tích hợp và cho phép nhà nghiên cứu có thể kiểm định nhiều giả thuyết khác nhau liên quan tới các phần tử vector. Số lượng mối quan hệ đồng tích hợp được tìm thấy sẽ tương ứng với số mối quan hệ dài hạn trong mô hình

Kiểm định đồng tích hợp của Johansen :

• $r = 0$:

$$\begin{cases} H_0 : r = 0, & \text{không có mối quan hệ đồng tích hợp} \\ H_1 : r > 0, & \text{có ít nhất 1 mối quan hệ đồng tích hợp} \end{cases}$$

• $r \leq 1$:

$$\begin{cases} H_0 : r \leq 1, & \text{có 1 mối quan hệ đồng tích hợp} \\ H_1 : r > 1, & \text{có nhiều hơn 1 mối quan hệ đồng tích hợp} \end{cases}$$

• $r \leq 2$:

$$\begin{cases} H_0 : r \leq 2, & \text{có 2 mối quan hệ đồng tích hợp} \\ H_1 : r > 2, & \text{có nhiều hơn 2 mối quan hệ đồng tích hợp} \end{cases}$$

• $r \leq 3$:

$$\begin{cases} H_0 : r \leq 3, & \text{có 3 mối quan hệ đồng tích hợp} \\ H_1 : r > 3, & \text{có nhiều hơn 3 mối quan hệ đồng tích hợp} \end{cases}$$

Sau khi thực hiện kiểm định ta thu được bảng kết quả

Bảng 4.10: Kiểm định Johansen giữa các chuỗi giá

| H_0 | Test | 10pct | 5pct | 1pct |
|------------|-------|-------|-------|--------|
| $r \leq 3$ | 0.00 | 6.50 | 8.18 | 11.65 |
| $r \leq 2$ | 6.91 | 15.6 | 17.95 | 23.52 |
| $r \leq 1$ | 16.42 | 28.71 | 31.52 | 37.222 |
| $r = 0$ | 33.01 | 45.23 | 48.28 | 53.43 |

Với mức ý nghĩa $\alpha = 5\%$, ta có $33.01 < 48.28$ nên chưa bác bỏ H_0 , số mối quan hệ đồng tích hợp bằng 0. Như vậy, **không có mối quan hệ đồng tích hợp nào giữa 4 chuỗi giá VCB, CTG, PVC, PVS.**

4.4 Phân tích VAR

Sau khi các bài cá nhân đã sử dụng mô hình đơn biến để dự báo chuỗi giá cổ phiếu, thì đến phần này, mô hình VAR được sử dụng với bốn chuỗi giá để đánh giá tác động qua lại giữa các chuỗi, cũng như đánh giá khả năng dự báo của VAR so với các mô hình đơn biến trên.

Mô hình hồi quy VAR đòi hỏi các chuỗi đầu vào phải dừng. Dựa trên kết quả kiểm định từ phần bài cá nhân, ta có chuỗi log-return (chuỗi lợi suất liên tục) của cả bốn mã cổ phiếu VCB, CTG, PVC, PVS đều dừng. Vì vậy, nhóm nghiên cứu sử dụng chuỗi log-return để dự báo cho mô hình VAR.

4.4.1 Xác định bậc của trẽ

| \$selection | | | |
|-------------|-------|-------|--------|
| AIC(n) | HQ(n) | SC(n) | FPE(n) |
| 1 | 1 | 1 | 1 |

Hình 4.8: Xác định bậc của trẽ

Kết luận: Tất cả các tiêu chí AIC, HQ, SC và FPE đều xác định mô hình có bậc của trẽ là 1. Vậy ta sẽ phân tích và dự báo trên mô hình VAR với trẽ bậc 1.

4.4.2 Ước lượng mô hình VAR(1)

Hồi quy mô hình VAR với $p = 1$, thu được kết quả như bảng 4.11, 4.12, 4.13, 4.14.

Bảng 4.11: Hồi quy VAR cho r_{VCB}

| | Estimate | Std. Error | P-value |
|---------------------|----------|------------|---------|
| rVCB.I1 | -0.05948 | 0.04768 | 0.213 |
| rCTG.I1 | 0.03083 | 0.03835 | 0.422 |
| rPVC.I1 | 0.03812 | 0.03593 | 0.289 |
| rPVS.I1 | -0.05630 | 0.04587 | 0.220 |
| Adjusted R-squared: | 0.001297 | | |

Ghi chú: ***, **, *; chỉ 0.1%, 1%, 5%, 10%

Mô hình VAR(1) cho r_{VCB} là:

$$r_{VCB_t} = -0.05948r_{VCB_{t-1}} + 0.03083r_{CTG_{t-1}} + 0.03812r_{PVC_{t-1}} - 0.05630r_{PVS_{t-1}}$$

Bảng 4.12: Hồi quy VAR cho r_{CTG}

| | Estimate | Std. Error | P-value |
|---------------------|----------|------------|----------|
| rVCB.I1 | -0.04300 | 0.06383 | 0.5009 |
| rCTG.I1 | -0.11531 | 0.05135 | 0.0252 * |
| rPVC.I1 | 0.01476 | 0.04810 | 0.7590 |
| rPVS.I1 | -0.05029 | 0.06140 | 0.4132 |
| Adjusted R-squared: | 0.01581 | | |

Ghi chú: ***, **, *; chỉ 0.1%, 1%, 5%, 10%

Mô hình VAR(1) cho r_{CTG} là:

$$r_{CTG_t} = -0.04300r_{VCB_{t-1}} - 0.11531r_{CTG_{t-1}} + 0.01476r_{PVC_{t-1}} - 0.05029r_{PVS_{t-1}}$$

Bảng 4.13: Hồi quy VAR cho r_{PVC}

| | Estimate | Std. Error | P-value |
|---------------------|----------|------------|---------|
| rVCB.I1 | 0.03020 | 0.09863 | 0.760 |
| rCTG.I1 | 0.04320 | 0.07934 | 0.586 |
| rPVC.I1 | -0.18977 | 0.07432 | 0.011 * |
| rPVS.I1 | 0.09387 | 0.09488 | 0.323 |
| Adjusted R-squared: | 0.009539 | | |

Ghi chú: ***, **, *; chỉ 0.1%, 1%, 5%, 10%

Mô hình VAR(1) cho r_{PVC} là:

$$r_{PVC_t} = 0.03020r_{VCB_{t-1}} + 0.04320r_{CTG_{t-1}} - 0.18977r_{PVC_{t-1}} + 0.09387r_{PVS_{t-1}}$$

Bảng 4.14: Hồi quy VAR cho r_{PVS}

| | Estimate | Std. Error | P-value |
|-----------------------------|----------|------------|---------|
| rVCB.I1 | 0.01119 | 0.08015 | 0.8890 |
| rCTG.I1 | 0.01072 | 0.06447 | 0.8680 |
| rPVC.I1 | -0.10352 | 0.06039 | 0.0871 |
| rPVS.I1 | -0.02537 | 0.07710 | 0.7423 |
| const | 0.089797 | 0.091119 | 0.3249 |
| Adjusted R-squared: 0.01323 | | | |

Ghi chú: ***, **, *; chỉ 0.1%, 1%, 5%, 10%

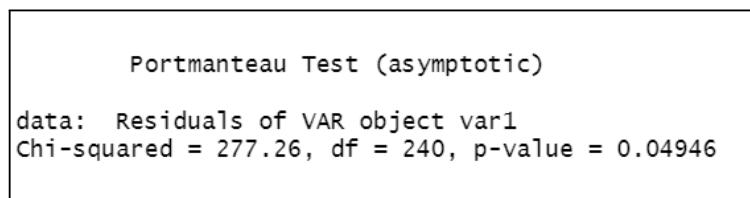
Mô hình VAR(1) cho r_{PVS} là:

$$r_{PVS_t} = 0.01119r_{VCB_{t-1}} + 0.01072r_{CTG_{t-1}} - 0.10352r_{PVC_{t-1}} - 0.02537r_{PVS_{t-1}}$$

4.4.3 Kiểm định tương quan phần dư

Kiểm định Portmanteau với cặp giả thuyết:

$$\begin{cases} H_0 : \text{Phần dư không có tự tương quan} \\ H_1 : \text{Phần dư có tự tương quan} \end{cases}$$

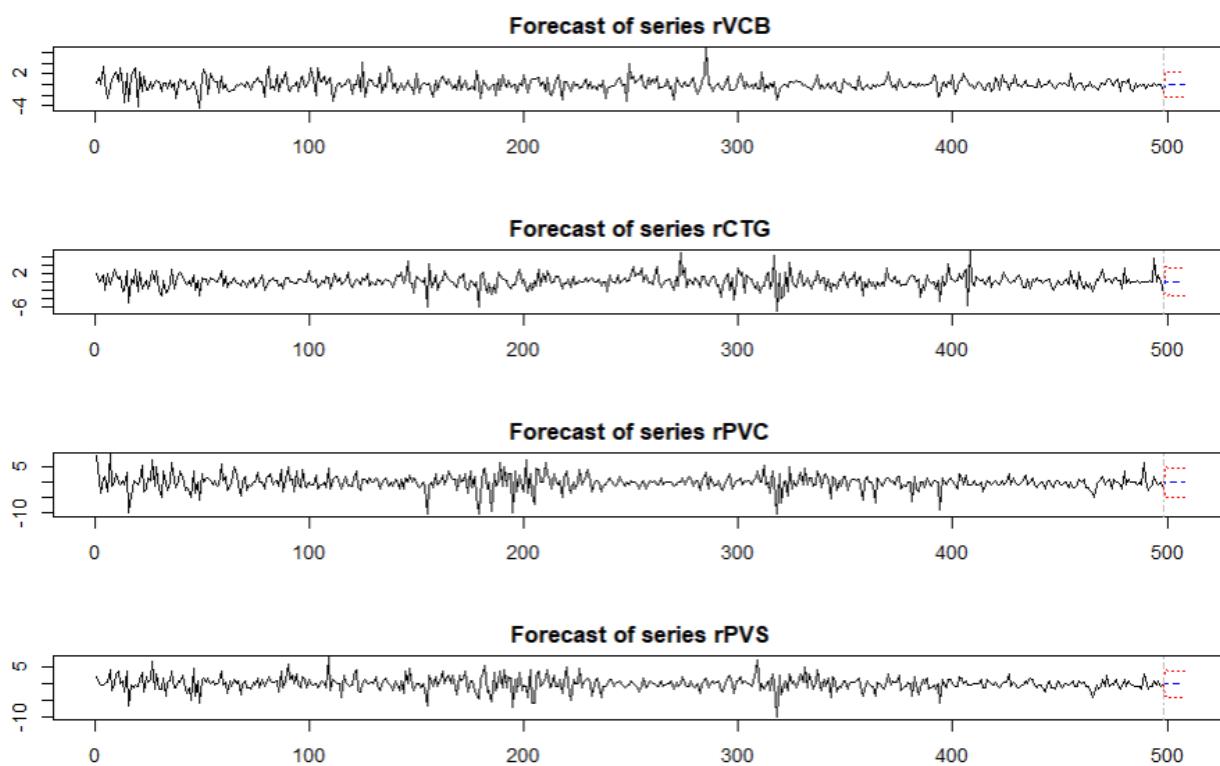


Hình 4.9: Kiểm định tương quan phần dư

Sau khi thực hiện kiểm định, thu được kết quả $p-value = 0.04946$. Với mức ý nghĩa 1%, chưa bác bỏ giả thiết H_0 , các phần dư không xảy ra hiện tượng tự tương quan, do đó mô hình VAR(1) là mô hình tốt để dự báo.

4.4.4 Dự báo bằng mô hình VAR(1)

Sử dụng lệnh ‘plot’ để biểu diễn đồ thị hàm dự báo:

Hình 4.10: Đồ thị dự báo r_{VCB} , r_{CTG} , r_{PVC} , r_{PVS}

So sánh kết quả dự báo của từng chuỗi cá nhân

Sử dụng kết quả dự báo từ mô hình ARIMA tốt nhất trong phần cá nhân, so sánh với kết quả dự báo từ mô hình VAR(1). Ta lựa chọn mô hình tốt nhất dựa trên tiêu chí RMSE nhỏ nhất.

Dưới đây là phần kết quả dự báo từ cả hai mô hình và kết luận với từng chuỗi giá đóng cửa.

| Phiên | Log-return dự báo (%) | Giá quy đổi từ log-return dự báo (VND) | Giá dự báo từ ARIMA (VND) | Giá thực (VND) |
|-------|-----------------------|--|---------------------------|----------------|
| 1 | -0.0732 | 91133.30 | 91311.7 | 91900 |
| 2 | 0.0133 | 91145.44 | 91413.2 | 92000 |
| 3 | -0.0013 | 91144.28 | 91596.6 | 92900 |
| 4 | 0.0001 | 91144.38 | 91626.2 | 92300 |
| 5 | 0.0000 | 91144.37 | 91662.3 | 92400 |
| 6 | 0.0000 | 91144.37 | 91696.4 | 91900 |
| 7 | 0.0000 | 91144.37 | 91749.3 | 91400 |
| 8 | 0.0000 | 91144.37 | 91800.3 | 91900 |
| 9 | 0.0000 | 91144.37 | 91850.7 | 91700 |
| 10 | 0.0000 | 91144.37 | 91899.0 | 91400 |
| RMSE | | 943.0836 | 751.8929 | |

Bảng 4.15: Dự báo giá cổ phiếu VCB 10 phiên đầu năm 2025

Như vậy, dựa vào kết quả dự báo của bảng 4.15, ta kết luận mô hình ARIMA cho kết quả dự báo giá đóng cửa của VCB tốt hơn.

| Phiên | Log-return dự báo (%) | Giá quy đổi từ log-return dự báo (VND) | Giá dự báo từ ARIMA (VND) | Giá thực (VND) |
|-------------|-----------------------|--|---------------------------|----------------|
| 1 | 0.3690 | 37939.74 | 37184.41 | 38,100.00 |
| 2 | -0.0449 | 37922.71 | 37184.68 | 37,200.00 |
| 3 | 0.0055 | 37924.79 | 37297.06 | 37,200.00 |
| 4 | -0.0006 | 37924.55 | 37358.98 | 37,450.00 |
| 5 | 0.0001 | 37924.57 | 37386.93 | 37,800.00 |
| 6 | 0.0000 | 37924.57 | 37414.91 | 37,400.00 |
| 7 | 0.0000 | 37924.57 | 37442.90 | 37,500.00 |
| 8 | 0.0000 | 37924.57 | 37470.91 | 37,700.00 |
| 9 | 0.0000 | 37924.57 | 37498.95 | 36,950.00 |
| 10 | 0.0000 | 37924.57 | 37527.01 | 36,750.00 |
| RMSE | | 644.1503 | 445.8601 | |

Bảng 4.16: Dự báo giá cổ phiếu CTG 10 phiên đầu năm 2025

Qua đó, ta thấy kết quả dự báo giá đóng cửa của CTG từ mô hình ARIMA tốt hơn mô hình VAR.

| Phiên | Log-return dự báo (%) | Giá quy đổi từ log-return dự báo (VND) | Giá dự báo từ ARIMA (VND) | Giá thực (VND) |
|-------------|-----------------------|--|---------------------------|----------------|
| 1 | 0.1737 | 10418.08 | 10438.5 | 10500 |
| 2 | 0.0042 | 10417.65 | 10444.52 | 10500 |
| 3 | -0.0025 | 10417.38 | 10425.96 | 10300 |
| 4 | 0.0007 | 10417.46 | 10442.79 | 10100 |
| 5 | -0.0001 | 10417.44 | 10405.63 | 10200 |
| 6 | 0.0000 | 10417.45 | 10437.46 | 10100 |
| 7 | 0.0000 | 10417.45 | 10419.43 | 10100 |
| 8 | 0.0000 | 10417.45 | 10426.49 | 10200 |
| 9 | 0.0000 | 10417.45 | 10426.09 | 10200 |
| 10 | 0.0000 | 10417.45 | 10420.18 | 10500 |
| RMSE | | 218.6917 | 225.2126 | |

Bảng 4.17: Dự báo giá cổ phiếu PVC 10 phiên đầu năm 2025

Riêng đối với giá cổ phiếu PVC, mô hình VAR(1) cho kết quả tốt hơn.

| Phiên | Log-return dự báo (%) | Giá quy đổi từ log-return dự báo (VND) | Giá dự báo từ ARIMA (VND) | Giá thực (VND) |
|-------------|-----------------------|--|---------------------------|----------------|
| 1 | 0.1605 | 33954.45 | 33950.63 | 33800 |
| 2 | -0.0189 | 33948.03 | 33906.15 | 34000 |
| 3 | 0.0006 | 33948.22 | 33971.67 | 33400 |
| 4 | 0.0003 | 33948.32 | 33921.67 | 32100 |
| 5 | -0.0001 | 33948.29 | 33939.53 | 32100 |
| 6 | 0.0000 | 33948.30 | 33930.27 | 32200 |
| 7 | 0.0000 | 33948.30 | 33935.07 | 32000 |
| 8 | 0.0000 | 33948.30 | 33932.58 | 32600 |
| 9 | 0.0000 | 33948.30 | 33933.87 | 32100 |
| 10 | 0.0000 | 33948.30 | 33933.20 | 31700 |
| RMSE | | 1558.876 | 1546.77 | |

Bảng 4.18: Dự báo giá cổ phiếu PVS 10 phiên đầu năm 2025

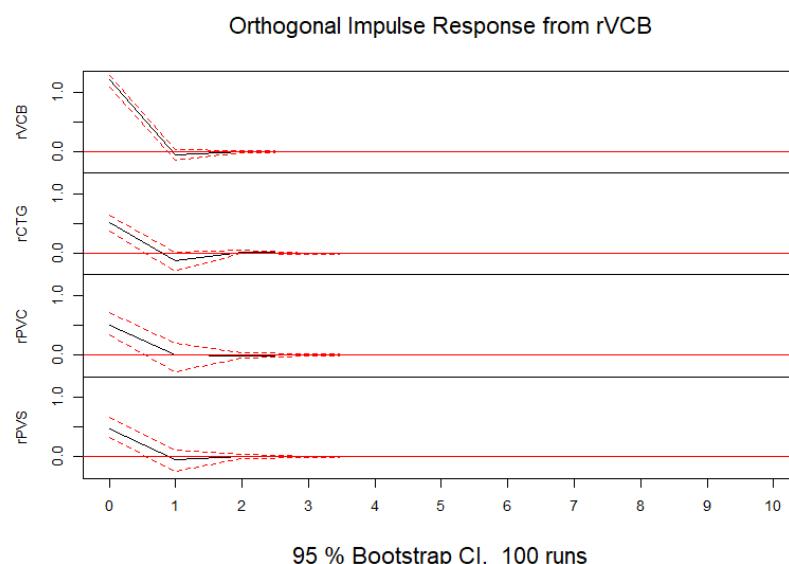
Từ kết quả bảng 4.18, ta thấy mô hình ARIMA(1,0,5) cho kết quả dự báo giá đóng cửa PVS tốt hơn mô hình VAR(1).

4.4.5 Phân tích hàm phản ứng và phân rã phương sai

Hàm phản ứng

Hàm phản ứng (Impulse response function) đo mức độ nhạy cảm của các chuỗi giá cổ phiếu với các cú sốc từ các chuỗi khác trong mô hình.

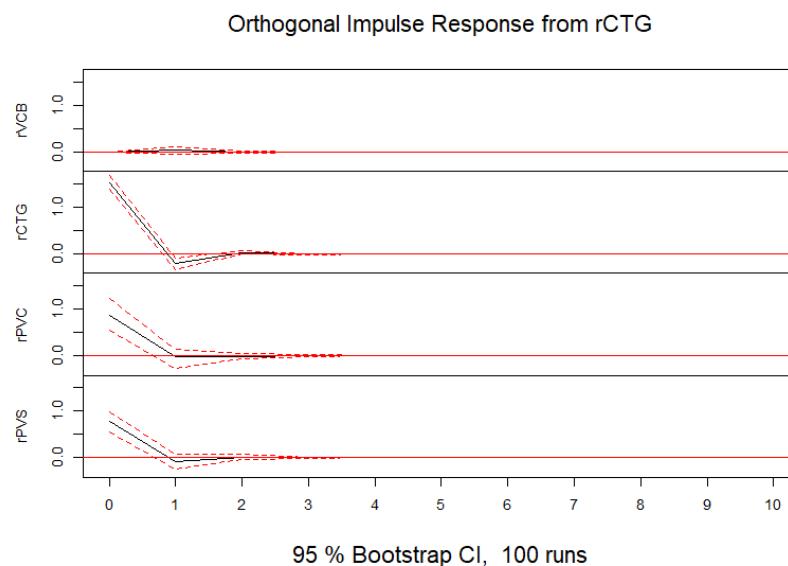
- Hàm phản ứng của r_{VCB}

Hình 4.11: Hàm phản ứng của r_{VCB}

Nhận xét: Đồ thị hàm phản ứng từ log-return của VCB cho thấy ảnh hưởng của cú sốc đến cả 4 chuỗi log-return đều có xu hướng giảm dần và tắt hẳn. Cụ thể, phản ứng của sự thay đổi r_{VCB} chỉ

sau 2 ngày là hết khi chịu cú sốc từ chính nó. Phản ứng của r_{CTG} và r_{PVS} cũng thể hiện xu hướng tương tự. Trong khi đó, phản ứng của r_{PVC} lại nhanh hơn, kết thúc hoàn toàn sau 1 ngày.

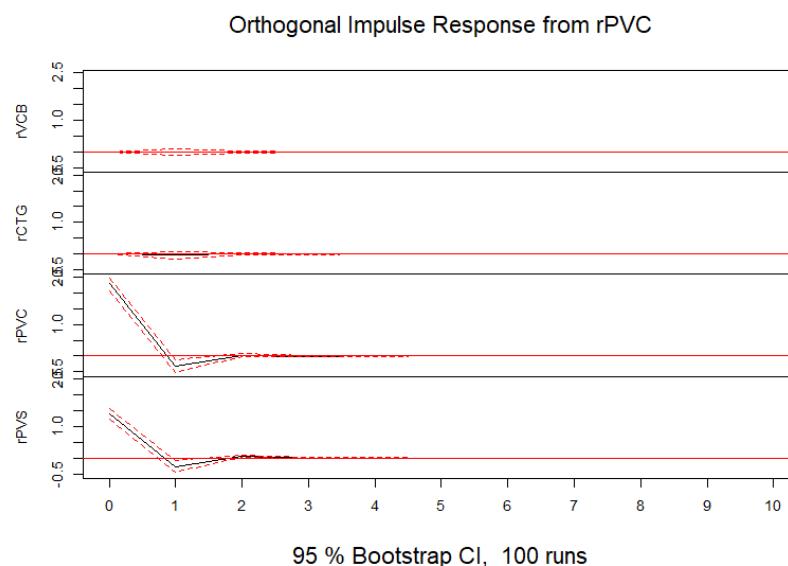
- Hàm phản ứng của r_{CTG}



Hình 4.12: Hàm phản ứng của r_{CTG}

Nhận xét: Từ đồ thị hàm phản ứng từ log-return của CTG, ảnh hưởng của cú sốc ở một thời điểm đến sự thay đổi log-return của chính nó giảm dần và kết thúc sau 2 ngày. Phản ứng của sự thay đổi r_{PVS} cũng cho thấy xu hướng tương tự. Trong khi đó, r_{VCB} lại không phản ứng với cú sốc từ r_{CTG} và phản ứng của sự thay đổi r_{PVC} cũng giảm dần, kết thúc sau 1 ngày.

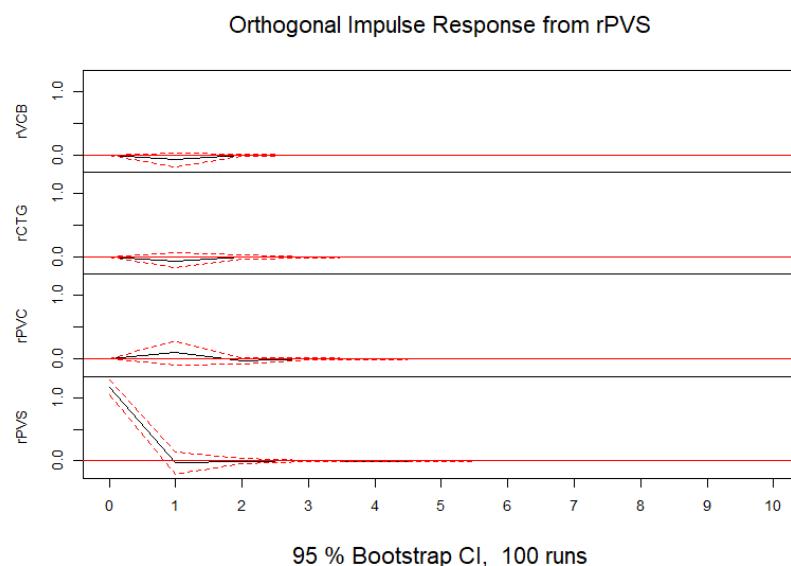
- Hàm phản ứng của r_{PVC}



Hình 4.13: Hàm phản ứng của r_{PVC}

Nhận xét: Với đồ thị hàm phản ứng từ log-return của PVC, ảnh hưởng của cú sốc ở một thời điểm đến sự thay đổi log-return của chính nó giảm dần và nhanh chóng kết thúc sau 2 ngày. Phản ứng của sự thay đổi r_{PVS} cũng cho thấy xu hướng tương tự. Tuy nhiên, r_{VCB} và r_{CTG} lại không phản ứng với cú sốc từ r_{PVC} .

- Hàm phản ứng của r_{PVS}

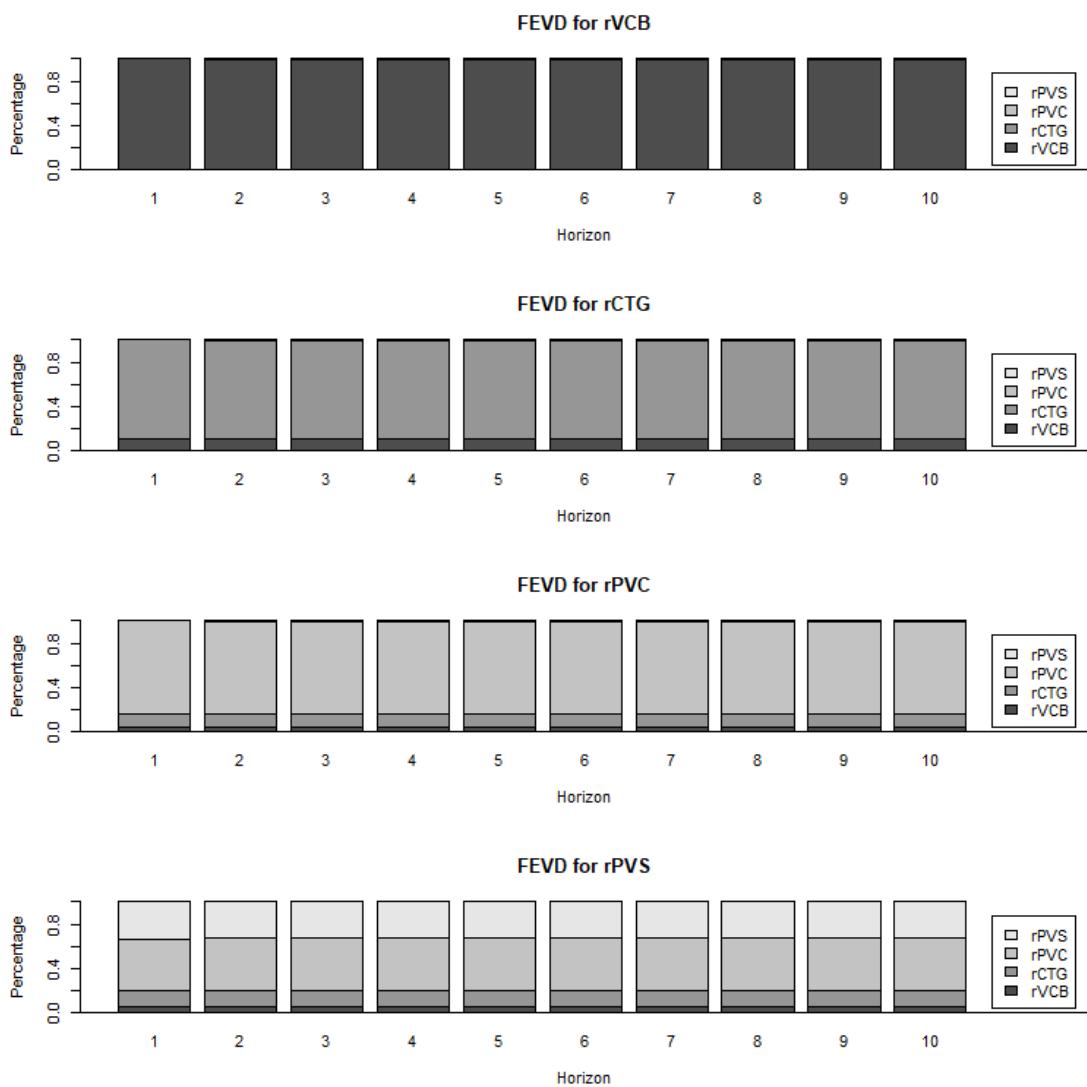


Hình 4.14: Hàm phản ứng của r_{PVS}

Nhận xét: Với đồ thị hàm phản ứng từ log-return của PVS, ta nhận thấy ảnh hưởng của cú sốc ở một thời điểm chỉ ảnh hưởng mạnh đến sự thay đổi log-return của chính nó mà gần như không ảnh hưởng đến 3 biến còn lại, đặc biệt r_{VCB}, r_{CTG} . Và tác động từ cú sốc tới chính nó cũng chỉ tồn tại trong khoảng 2 ngày sau với biên độ dao động giảm dần. Với r_{PVC} , cú sốc của r_{PVS} gây ra ảnh hưởng ở mức độ yếu trong 2 ngày với xu hướng khác nhau.

Hàm phân rã phương sai

Phân rã phương sai của sai số dự báo đo lường tỷ lệ (%) biến động của một biến trong mô hình VAR được giải thích bởi các cú sốc từ chính nó và từ các biến khác theo thời gian .



Hình 4.15: Hàm phân ra phương sai của 4 chuỗi log-return

Dựa vào đồ thị hình 4.15, ta có một số nhận xét sau:

- Sự biến động của r_{VCB} chủ yếu được giải thích bởi chính nó trong suốt 10 ngày, hầu như không bị ảnh hưởng từ sự biến động của r_{CTG} , r_{PVC} và r_{PVS} .
- Với r_{CTG} , sự biến động của nó cũng được giải thích phần lớn bởi chính nó. Cụ thể là 89.44% trong ngày đầu tiên và khoảng 88.87% trong 9 ngày tiếp theo. Phần còn lại được giải thích bởi r_{VCB} (10.55%-10.94%).
- Trong suốt 10 ngày, biến động của r_{PVC} được giải thích chủ yếu bởi chính nó, với mức độ lên tới 83.97%. Các yếu tố còn lại đóng góp vào sự biến động bao gồm r_{CTG} , chiếm khoảng từ 11.72% đến 11.91%, và r_{VCB} , với mức đóng góp dao động từ 4.06% đến 4.13%.
- Trong ngày đầu tiên, sự biến động của r_{PVS} chỉ phụ thuộc chính nó 33.2%, tăng dần lên và duy trì 33.54% trong 9 ngày tiếp theo. Phần lớn sự biến động của nó chịu tác động bởi r_{PVC} với 46.94% trong ngày đầu tiên và 47.76% trong giai đoạn tiếp. Phần còn lại được giải thích bởi r_{CTG} và r_{VCB} lần lượt với 14.44% và 5.4% trong suốt 10 ngày.

Qua đó, ta có thể thấy VCB là cổ phiếu có tính ổn định cao, ít chịu tác động từ biến động thị trường ngành khác. Bên cạnh đó, mã CTG phụ thuộc một phần vào biến động của VCB.

Hai mã cổ phiếu PVS, PVC cùng thuộc nhóm ngành dầu khí, có tác động lẫn nhau và ít bị ảnh hưởng bởi nhóm ngành ngân hàng.

Chương 5

KẾT LUẬN

Trong khuôn khổ đề tài "Phân tích và dự báo chuỗi giá cổ phiếu và chuỗi doanh thu", nhóm đã thực hiện nghiên cứu định lượng trên 4 doanh nghiệp tiêu biểu thuộc hai ngành then chốt: ngân hàng (VCB, CTG) và dầu khí (PVC, PVS). Từ kết quả phân tích, nhóm rút ra một số kết luận quan trọng sau:

Về dự báo doanh thu, nhóm đã thử nghiệm và so sánh nhiều mô hình như lin-lin, lin-log, log-lin, log-log, Holt-Winters để tìm ra mô hình phù hợp nhất cho từng công ty. Kết quả dự báo doanh thu cho năm 2025 của từng công ty đã được trình bày trong phần bài cá nhân.

Đối với dự báo giá cổ phiếu, các mô hình ARIMA và VAR đã được sử dụng và so sánh. Trong phần bài cá nhân, mô hình ARIMA tốt nhất cho từng cổ phiếu (VCB(3,0,0), CTG(0,0,4), PVC(5,1,1), PVS(1,0,5) đã được xác định để dự báo giá 10 phiên đầu năm 2025. Nhìn chung, phương pháp dự báo gián tiếp thông qua chuỗi log-return thường cho kết quả tốt hơn so với dự báo trực tiếp chuỗi giá đóng cửa.

Trong phần xây dựng danh mục tối ưu, hai phương án đã được xem xét: tối thiểu hóa phương sai và tối đa hóa lợi nhuận kỳ vọng. Phương án tối thiểu hóa phương sai cho thấy tỷ trọng tối ưu là 66.55% cho VCB, 22.04% cho CTG, 0% cho PVC, và 11.42% cho PVS, với phương sai tối thiểu là 1.161358 và lợi suất kỳ vọng là 6.4%. Phương án tối đa hóa lợi nhuận (sau khi điều chỉnh điều kiện về phương sai) cho thấy tỷ trọng tối ưu là 0% cho VCB, 21.296% cho CTG, 0% cho PVC, và 78.704% cho PVS, với lợi suất kỳ vọng cao nhất là 0.088981 và phương sai là 3.170958.

Dự báo chuỗi lợi suất danh mục đã được thực hiện theo hai cách, và kết quả cho thấy cách tính chuỗi lợi suất portfolio bằng cách kết hợp lợi suất dự báo cá nhân cho kết quả gần với dữ liệu thực tế hơn. Mô hình ARIMA(1,0,0) được chọn để dự báo chuỗi lợi suất danh mục.

Phân tích đồng tích hợp theo kiểm định Johansen cho thấy không có mối quan hệ đồng tích hợp nào giữa bốn chuỗi giá VCB, CTG, PVC, và PVS.

Mô hình VAR(1) đã được sử dụng để phân tích tác động qua lại và khả năng dự báo giữa các chuỗi log-return. So sánh với mô hình ARIMA tốt nhất của từng cổ phiếu, mô hình ARIMA cá nhân cho kết quả dự báo giá cổ phiếu VCB, CTG và PVS tốt hơn mô hình VAR(1). Phân tích hàm phản ứng cho thấy ảnh hưởng của các cú sốc có xu hướng giảm dần sau khoảng 1-2 ngày. Phân rã phương sai cho thấy sự biến động của rVCB và rCTG chủ yếu do chính nó giải thích, trong khi rPVC, rPVS chịu tác động lẫn nhau và ít chịu tác động từ nhóm ngành còn lại

ĐÁNH GIÁ

| Họ và tên | MSSV | Đóng góp | % đóng góp |
|--------------------|----------|---|------------|
| Phạm Hạnh Nguyên | 11224830 | Đồng tích hợp Xây dựng danh mục tối ưu Dự báo chuỗi portfolio từ giá quá khứ Viết báo cáo: Xây dựng danh mục tối ưu, dự báo chuỗi portfolio từ giá quá khứ | 25% |
| Trần Thị Kiều Oanh | 11225075 | Đồng tích hợp Viết báo cáo: Tổng quan, đồng tích hợp, kết luận | 25% |
| Nghiêm Gia Phương | 11225223 | Xây dựng danh mục tối ưu Dự báo chuỗi portfolio từ chuỗi giá dự báo, so sánh 2 cách Mô hình VAR Viết báo cáo: Mở đầu, VAR | 25% |
| Nguyễn Thị Trà | 11226303 | Xây dựng danh mục tối ưu Dự báo chuỗi portfolio từ chuỗi giá dự báo, so sánh 2 cách Mô hình VAR Viết báo cáo: Tổng quan, dự báo chuỗi portfolio từ chuỗi dự báo, so sánh 2 cách, VAR | 25% |

PHỤ LỤC

Bài cá nhân - Phạm Hạnh Nguyên

Mô hình lin - lin:

```
## 
## Call:
## lm(formula = Dt_trainset ~ time_trainset)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -2973.7 -1946.5  -855.8  1415.2  6745.4 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 2471.94    668.60   3.697 0.000511 ***
## time_trainset 363.05     20.41  17.791 < 2e-16 ***
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 2468 on 54 degrees of freedom
## Multiple R-squared:  0.8543, Adjusted R-squared:  0.8516 
## F-statistic: 316.5 on 1 and 54 DF,  p-value: < 2.2e-16
```

Bảng 5.1: Kết quả hồi quy lin - lin

Mô hình lin - log:

```

## 
## Call:
## lm(formula = Dt_trainset ~ log(time_trainset))
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5404.7 -3664.9   451.5  2100.9 11127.0
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)            -3650.5     2064.5  -1.768   0.0827 .  
## log(time_trainset)    5351.2      644.5   8.303 3.21e-11 *** 
## ---                
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 4285 on 54 degrees of freedom
## Multiple R-squared:  0.5608, Adjusted R-squared:  0.5526 
## F-statistic: 68.94 on 1 and 54 DF,  p-value: 3.207e-11

```

Bảng 5.2: Kết quả hồi quy lin - log

Mô hình log - lin:

```

## 
## Call:
## lm(formula = log(Dt_trainset) ~ time_trainset)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.23517 -0.10908 -0.01643  0.10467  0.41759
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)        8.530198   0.042284 201.74 <2e-16 *** 
## time_trainset 0.028440   0.001291  22.04 <2e-16 *** 
## ---                
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.1561 on 54 degrees of freedom
## Multiple R-squared:  0.8999, Adjusted R-squared:  0.8981 
## F-statistic: 485.7 on 1 and 54 DF,  p-value: < 2.2e-16

```

Bảng 5.3: Kết quả hồi quy log - lin

Mô hình log - log:

```

## 
## Call:
## lm(formula = log(Dt_trainset) ~ log(time_trainset))
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.42608 -0.23163  0.07427  0.17722  0.51383
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)            7.94136   0.13130  60.48 < 2e-16 ***
## log(time_trainset)    0.45468   0.04099  11.09 1.53e-15 ***
## ---                
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.2725 on 54 degrees of freedom
## Multiple R-squared:  0.695, Adjusted R-squared:  0.6894 
## F-statistic: 123.1 on 1 and 54 DF,  p-value: 1.53e-15

```

Bảng 5.4: Kết quả hồi quy log - log

Hồi quy với biến giả mùa vụ:

```

## 
## Call:
## lm(formula = Dt_trainset ~ s2 + s3 + s4)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8165  -5055  -2359   4260  16146
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)  12356.8     1757.9   7.029 4.44e-09 ***
## s2          319.3      2486.0   0.128   0.898    
## s3          506.9      2486.0   0.204   0.839    
## s4         1022.3     2486.0   0.411   0.683    
## ---                
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 6577 on 52 degrees of freedom
## Multiple R-squared:  0.00341, Adjusted R-squared:  -0.05409 
## F-statistic: 0.0593 on 3 and 52 DF,  p-value: 0.9809

```

Bảng 5.5: Kết quả hồi quy với biến giả mùa vụ

Xu thế tuyến tính + Mùa vụ, dạng cộng:

```

## 
## Call:
## lm(formula = Dt_trainset ~ time_trainset + s2 + s3 + s4)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3011.3 -1981.9 - 749.2 1554.5 6702.0
## 
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 2549.49    884.78   2.882 0.00578 **  
## time_trainset 363.23    21.04  17.267 < 2e-16 ***  
## s2          -43.89    959.63  -0.046 0.96370    
## s3          -219.54    960.32  -0.229 0.82009    
## s4          -67.40    961.48  -0.070 0.94438    
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 2538 on 51 degrees of freedom
## Multiple R-squared:  0.8544, Adjusted R-squared:  0.843  
## F-statistic: 74.84 on 4 and 51 DF,  p-value: < 2.2e-16

```

Bảng 5.6: Kết quả xu thế tuyến tính có mùa vụ, dạng cộng

Xu thế tuyến tính + Mùa vụ, dạng nhân:

```

## 
## Call:
## lm(formula = Dt_trainset ~ time_trainset + I(time_trainset * 
##           s2) + I(time_trainset * s3) + I(time_trainset * s4))
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3037.1 -1970.6 - 808.8 1460.8 6587.0
## 
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 2470.0313   688.1366   3.589 0.000743 ***  
## time_trainset 363.4982   28.6081  12.706 < 2e-16 ***  
## I(time_trainset * s2)  2.5235   30.1095   0.084 0.933536    
## I(time_trainset * s3) -3.4846   29.7335  -0.117 0.907166    
## I(time_trainset * s4) -0.4262   29.3829  -0.015 0.988484    
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 2539 on 51 degrees of freedom
## Multiple R-squared:  0.8544, Adjusted R-squared:  0.843  
## F-statistic: 74.81 on 4 and 51 DF,  p-value: < 2.2e-16

```

Bảng 5.7: Kết quả xu thế tuyến tính có mùa vụ, dạng nhân

Holt - Winter có mùa vụ, dạng cộng:

```

## Holt-Winters exponential smoothing with trend and additive seasonal
## component.
##
## Call:
## HoltWinters(x = Dt_trainset, seasonal = "a")
##
## Smoothing parameters:
## alpha: 0.700483
## beta : 0.9121006
## gamma: 1
##
## Coefficients:
##                 [,1]
## a    25491.96811
## b    -2113.55510
## s1   270.31276
## s2   -32.47053
## s3  -1209.92944
## s4  -394.17711

```

Bảng 5.8: Kết quả Holt - Winter có mùa vụ, dạng cộng

Holt - Winter có mùa vụ, dạng nhân:

```

## Warning in HoltWinters(Dt_trainset, seasonal = "m") : optimization
## difficulties:
## ERROR: ABNORMAL_TERMINATION_IN_LNSRCH
reg9
##
## Holt-Winters exponential smoothing with trend and multiplicative
## seasonal component.
##
## Call:
## HoltWinters(x = Dt_trainset, seasonal = "m")
##
## Smoothing parameters:
## alpha: 0.8639038
## beta : 0.07387855
## gamma: 1
##
## Coefficients:
##                 [,1]
## a    2.644219e+04
## b    5.334387e+02
## s1   9.841626e-01
## s2   9.870878e-01
## s3   9.524046e-01
## s4   9.491571e-01

```

Bảng 5.9: Kết quả Holt - Winter có mùa vụ, dạng nhân

Holt - Winter có mùa vụ, dạng cộng; sử dụng để dự báo 2025:

```
## Holt-Winters exponential smoothing with trend and additive seasonal component.  
##  
## Call:  
## HoltWinters(x = DT, seasonal = "a")  
##  
## Smoothing parameters:  
## alpha: 0.8934877  
## beta : 0.5406154  
## gamma: 1  
##  
## Coefficients:  
## [,1]  
## a 23716.90698  
## b 89.99538  
## s1 59.43605  
## s2 84.07099  
## s3 -164.61405  
## s4 -135.64698
```

Bảng 5.10: Holt - Winter có mùa vụ, dạng cộng; sử dụng để dự báo cho 2025

Kiểm định ADF cho chuỗi giá của VCB, có xu thế:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression trend
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
## 
## Residuals:
##      Min        1Q    Median        3Q       Max
## -3677.1   -562.0   -15.1    511.6   6444.8
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 3293.05627  956.97527  3.441 0.000629 ***
## z.lag.1     -0.04009   0.01205 -3.326 0.000946 *** 
## tt          0.94265   0.49445  1.906 0.057173 .  
## z.diff.lag  -0.03379   0.04483 -0.754 0.451327    
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 1013 on 493 degrees of freedom
## Multiple R-squared:  0.02626,   Adjusted R-squared:  0.02033 
## F-statistic: 4.432 on 3 and 493 DF,  p-value: 0.004359
## 
## 
## Value of test-statistic is: -3.3261 4.3474 6.0396
## 
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.98 -3.42 -3.13
## phi2  6.15  4.71  4.05
## phi3  8.34  6.30  5.36

```

Bảng 5.11: Kiểm định ADF cho chuỗi giá của VCB, có xu thế

Kiểm định ADF cho chuỗi sai phân của VCB, có xu thế:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression trend
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -3516.6  -558.1  -10.6   504.1  6380.1 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 122.84515  92.63045   1.326   0.185    
## z.lag.1     -1.10163   0.06529 -16.872 <2e-16 ***  
## tt          -0.31190   0.32147  -0.970   0.332    
## z.diff.lag    0.04825   0.04505   1.071   0.285    
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 1023 on 492 degrees of freedom
## Multiple R-squared:  0.5264, Adjusted R-squared:  0.5235 
## F-statistic: 182.3 on 3 and 492 DF,  p-value: < 2.2e-16 
## 
## 
## Value of test-statistic is: -16.8717 94.8877 142.3271
## 
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.98 -3.42 -3.13
## phi2  6.15  4.71  4.05
## phi3  8.34  6.30  5.36

```

Bảng 5.12: Kiểm định ADF cho chuỗi sai phân của VCB, có xu thế

Kiểm định ADF cho chuỗi lợi suất của VCB, có xu thế:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression trend
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -4.7731 -0.6388 -0.0093  0.5784  6.7843 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 0.1570320 0.1088796 1.442   0.150    
## z.lag.1    -1.1100329 0.0654363 -16.964 <2e-16 ***  
## tt        -0.0004040 0.0003778 -1.069   0.285    
## z.diff.lag  0.0502420 0.0450154  1.116   0.265    
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 1.202 on 492 degrees of freedom
## Multiple R-squared:  0.5298, Adjusted R-squared:  0.5269 
## F-statistic: 184.8 on 3 and 492 DF,  p-value: < 2.2e-16 
## 
## 
## Value of test-statistic is: -16.9636 95.9243 143.8813 
## 
## Critical values for test statistics:
##      1pct  5pct 10pct 
## tau3 -3.98 -3.42 -3.13 
## phi2  6.15  4.71  4.05 
## phi3  8.34  6.30  5.36

```

Bảng 5.13: Kiểm định ADF cho chuỗi lợi suất của VCB, có xu thế

Kiểm định ADF cho chuỗi giá của VCB, có hệ số chặn:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression drift
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3679.5  -539.5   -15.7   513.3  6419.0
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 1.996e+03 6.749e+02  2.958 0.00325 **  
## z.lag.1     -2.246e-02 7.748e-03 -2.898 0.00392 **  
## z.diff.lag  -4.344e-02 4.466e-02 -0.973 0.33123  
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 1016 on 494 degrees of freedom
## Multiple R-squared:  0.01908,    Adjusted R-squared:  0.01511 
## F-statistic: 4.805 on 2 and 494 DF,  p-value: 0.008578 
## 
## 
## Value of test-statistic is: -2.8982 4.6789
## 
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phi1  6.47  4.61  3.79

```

Bảng 5.14: Kiểm định ADF cho chuỗi giá của VCB, có hệ số chặn

Kiểm định ADF cho chuỗi sai phân của VCB, có hệ số chặn:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression drift
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -3449.9  -570.9   -33.1   518.7  6365.5 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 44.86295  46.04407  0.974   0.330    
## z.lag.1     -1.09801  0.06518 -16.845  <2e-16 ***  
## z.diff.lag    0.04643  0.04500   1.032   0.303    
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 1023 on 493 degrees of freedom
## Multiple R-squared:  0.5255, Adjusted R-squared:  0.5236 
## F-statistic: 273 on 2 and 493 DF,  p-value: < 2.2e-16 
## 
## 
## Value of test-statistic is: -16.8448 141.8778
## 
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phi1  6.47  4.61  3.79

```

Bảng 5.15: Kiểm định ADF cho chuỗi sai phân của VCB, có hệ số chặn

Kiểm định ADF cho chuỗi lợi suất của VCB, có hệ số chặn:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression drift
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -4.6855 -0.6509 -0.0437  0.5869  6.7651 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 0.05599   0.05411   1.035   0.301    
## z.lag.1    -1.10555   0.06531 -16.927 <2e-16 ***  
## z.diff.lag  0.04799   0.04497   1.067   0.286    
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 1.202 on 493 degrees of freedom
## Multiple R-squared:  0.5287, Adjusted R-squared:  0.5268 
## F-statistic: 276.5 on 2 and 493 DF,  p-value: < 2.2e-16 
## 
## 
## Value of test-statistic is: -16.9274 143.2729
## 
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phi1  6.47  4.61  3.79

```

Bảng 5.16: Kiểm định ADF cho chuỗi lợi suất của VCB, có hệ số chặn

Kiểm định ADF cho chuỗi giá của VCB, không có hệ số chặn:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression none
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -3415.4   -562.2   -35.7    526.7   6340.8 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## z.lag.1     0.0004088  0.0005277   0.775   0.439    
## z.diff.lag -0.0487798  0.0449700  -1.085   0.279    
## 
## Residual standard error: 1024 on 495 degrees of freedom
## Multiple R-squared:  0.003426,   Adjusted R-squared:  -0.0006005 
## F-statistic: 0.8509 on 2 and 495 DF,   p-value: 0.4277  
## 
## 
## Value of test-statistic is: 0.7747
## 
## Critical values for test statistics:
##      1pct  5pct 10pct 
## tau1 -2.58 -1.95 -1.62

```

Bảng 5.17: Kiểm định ADF cho chuỗi giá của VCB, không có hệ số chặn

Kiểm định ADF cho chuỗi sai phân của VCB, không có hệ số chặn:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression none
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -3401.0   -525.1    13.0   558.5  6406.0 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## z.lag.1     -1.09398   0.06505 -16.818 <2e-16 ***
## z.diff.lag   0.04441   0.04495   0.988   0.324    
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 1023 on 494 degrees of freedom
## Multiple R-squared:  0.5246, Adjusted R-squared:  0.5227 
## F-statistic: 272.6 on 2 and 494 DF,  p-value: < 2.2e-16 
## 
## 
## Value of test-statistic is: -16.8177
## 
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62

```

Bảng 5.18: Kiểm định ADF cho chuỗi sai phân của VCB, không có hệ số chẵn

Kiểm định ADF cho chuỗi lợi suất của VCB, không có hệ số chẵn:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression none
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -4.6235 -0.5939  0.0140  0.6424  6.8156 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## z.lag.1    -1.10102   0.06517 -16.895 <2e-16 ***
## z.diff.lag  0.04572   0.04492   1.018   0.309    
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 1.202 on 494 degrees of freedom
## Multiple R-squared:  0.5277, Adjusted R-squared:  0.5258 
## F-statistic: 275.9 on 2 and 494 DF,  p-value: < 2.2e-16
## 
## 
## Value of test-statistic is: -16.8948
## 
## Critical values for test statistics:
##      1pct  5pct 10pct 
## tau1 -2.58 -1.95 -1.62

```

Bảng 5.19: Kiểm định ADF cho chuỗi lợi suất của VCB, không có hệ số chẵn

Mô hình ARIMA(3,1,3) cho chuỗi sai phân VCB:

```

## Series: price
## ARIMA(3,1,3)
## 
## Coefficients:
##          ar1      ar2      ar3      ma1      ma2      ma3  
##        0.3845 -0.8776  0.3298 -0.4366  0.8459 -0.4347 
## s.e.  0.3374  0.1035  0.3099  0.3237  0.1013  0.2860 
## 
## sigma^2 = 1033024: log likelihood = -4151.84
## AIC=8317.68   AICC=8317.91   BIC=8347.15
## 
## Training set error measures:
##           ME      RMSE      MAE      MPE      MAPE      MASE  
## Training set 51.32382 1009.224 732.6535 0.0558885 0.854253 0.9905964
##           ACF1
## Training set -0.004931871

```

Bảng 5.20: Mô hình ARIMA(3,1,3) cho chuỗi sai phân VCB

Mô hình ARIMA(1,1,3) cho chuỗi sai phân VCB:

```

## Series: price
## ARIMA(1,1,3)
##
## Coefficients:
##             ar1      ma1      ma2      ma3
##           -0.1813  0.1295 -0.0506 -0.1169
## s.e.    0.3571  0.3536  0.0502  0.0436
##
## sigma^2 = 1034972: log likelihood = -4153.26
## AIC=8316.53   AICC=8316.65   BIC=8337.58
##
## Training set error measures:
##                         ME      RMSE      MAE      MPE      MAPE
MASE
## Training set 52.86363 1012.226 737.4233 0.05759337 0.8593826
0.9970456
##                         ACF1
## Training set -0.003414901

```

Bảng 5.21: Mô hình ARIMA(1,1,3) cho chuỗi sai phân VCB

Mô hình ARIMA(3,1,1) cho chuỗi sai phân VCB:

```

## Series: price
## ARIMA(3,1,1)
##
## Coefficients:
##             ar1      ar2      ar3      ma1
##           -0.1599 -0.0554 -0.1256  0.1071
## s.e.    0.3213  0.0475  0.0457  0.3232
##
## sigma^2 = 1033776: log likelihood = -4152.98
## AIC=8315.96   AICC=8316.08   BIC=8337.01
##
## Training set error measures:
##                         ME      RMSE      MAE      MPE      MAPE
MASE
## Training set 52.11235 1011.641 737.0611 0.05674961 0.8589569
0.9965559
##                         ACF1
## Training set -0.002478227

```

Bảng 5.22: Mô hình ARIMA(3,1,1) cho chuỗi sai phân VCB

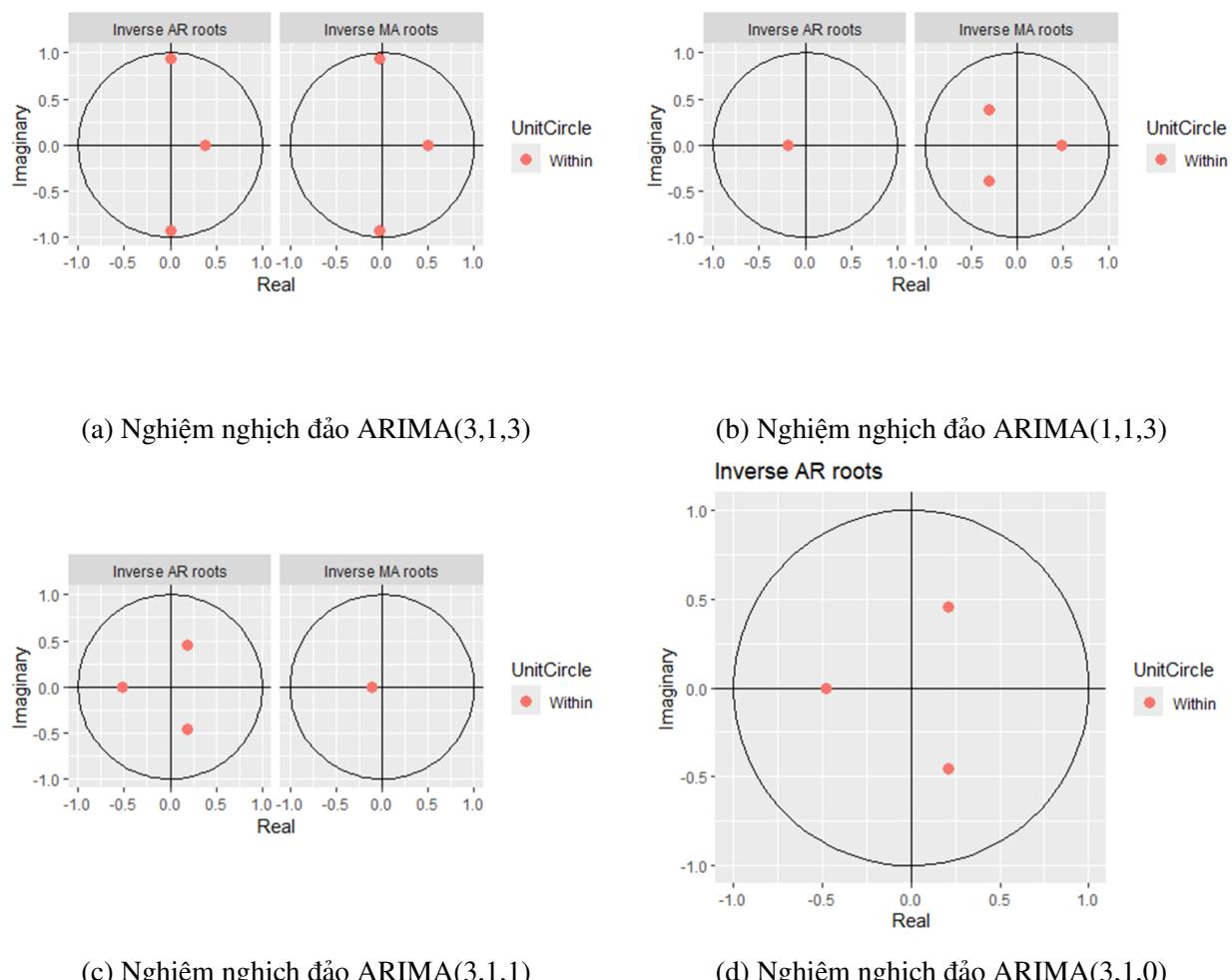
Mô hình ARIMA(3,1,0) cho chuỗi sai phân VCB:

```

## Series: price
## ARIMA(3,1,0)
##
## Coefficients:
##             ar1      ar2      ar3
##            -0.0545 -0.0503 -0.1204
## s.e.    0.0445  0.0445  0.0444
## 
## sigma^2 = 1031967: log likelihood = -4153.05
## AIC=8314.09   AICc=8314.17   BIC=8330.94
## 
## Training set error measures:
##               ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 52.74123 1011.778 737.66 0.05745949 0.8596644 0.9973655
## 
## ACF1
## Training set -0.0006199277

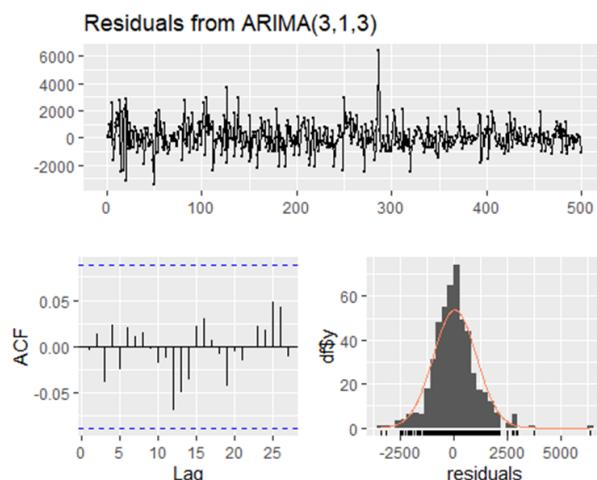
```

Bảng 5.23: Mô hình ARIMA(3,1,0) cho chuỗi sai phân VCB

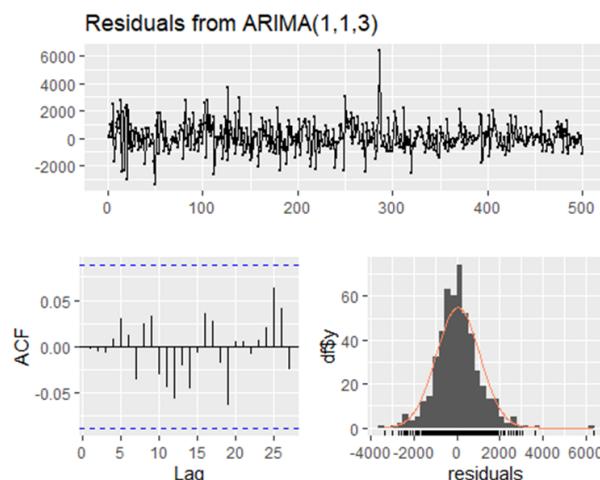
Nghiệm nghịch đảo đều nằm trong vòng tròn đơn vị:

Hình 5.1: Xem tính dừng qua nghiệm nghịch đảo

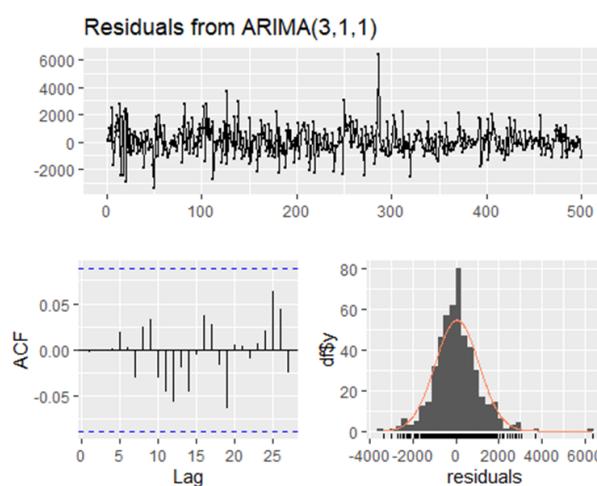
Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng:



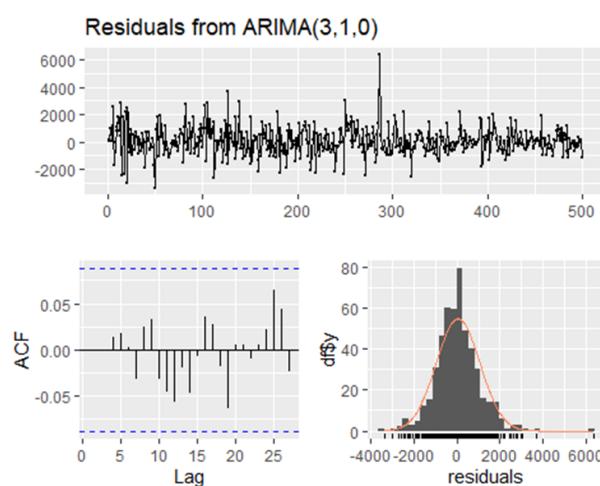
(a) Phân dư từ ARIMA(3,1,3)



(b) Phân dư từ ARIMA(1,1,3)



(c) Phân dư từ ARIMA(3,1,1)



(d) Phân dư từ ARIMA(3,1,0)

Hình 5.2: Kiểm định tính nhiễu trắng của phân dư

Mô hình ARIMA(3,0,3) cho chuỗi lợi suất VCB:

```
## Series: logreturn[2:499]
## ARIMA(3,0,3) with non-zero mean
##
## Coefficients:
## Warning in sqrt(diag(x$var.coef)) : NaNs produced
##          ar1      ar2      ar3      ma1      ma2      ma3    mean
##        0.0717 -0.6729  0.0394 -0.1326  0.6315 -0.1702  0.0556
##  s.e.      NaN       NaN       NaN       NaN       NaN     0.0815  0.0453
##
## sigma^2 = 1.43: log likelihood = -792.28
## AIC=1600.57  AICc=1600.86  BIC=1634.25
##
## Training set error measures:
##                         ME      RMSE      MAE      MPE      MAPE      MASE
ACF1
## Training set -0.00227107 1.187575 0.8579973 NaN Inf 0.6861076
-0.0005921046
```

Bảng 5.24: Mô hình ARIMA(3,0,3) cho chuỗi lợi suất VCB

Mô hình ARIMA(1,0,3) cho chuỗi lợi suất VCB:

```

## Series: logreturn[2:499]
## ARIMA(1,0,3) with non-zero mean
##
## Coefficients:
##             ar1      ma1      ma2      ma3      mean
##            -0.2308  0.1714 -0.0583 -0.1218  0.0532
## s.e.      0.3531  0.3495  0.0516  0.0434  0.0430
##
## sigma^2 = 1.428: log likelihood = -792.94
## AIC=1597.88   AICc=1598.05   BIC=1623.15
##
## Training set error measures:
##                         ME      RMSE      MAE      MPE      MAPE      MASE
ACF1
## Training set 0.0006214903 1.189181 0.8638282 NaN Inf 0.6907704
-0.0007601866

```

Bảng 5.25: Mô hình ARIMA(1,0,3) cho chuỗi lợi suất VCB

Mô hình ARIMA(3,0,1) cho chuỗi lợi suất VCB:

```

## Series: logreturn[2:499]
## ARIMA(3,0,1) with non-zero mean
##
## Coefficients:
##             ar1      ar2      ar3      ma1      mean
##            -0.2157 -0.0637 -0.1312  0.1551  0.0534
## s.e.      0.3110  0.0485  0.0453  0.3130  0.0436
##
## sigma^2 = 1.427: log likelihood = -792.64
## AIC=1597.27   AICc=1597.44   BIC=1622.54
##
## Training set error measures:
##                         ME      RMSE      MAE      MPE      MAPE      MASE
ACF1
## Training set 0.0003689131 1.188444 0.8635011 NaN Inf 0.6905088
0.0002222302

```

Bảng 5.26: Mô hình ARIMA(3,0,1) cho chuỗi lợi suất VCB

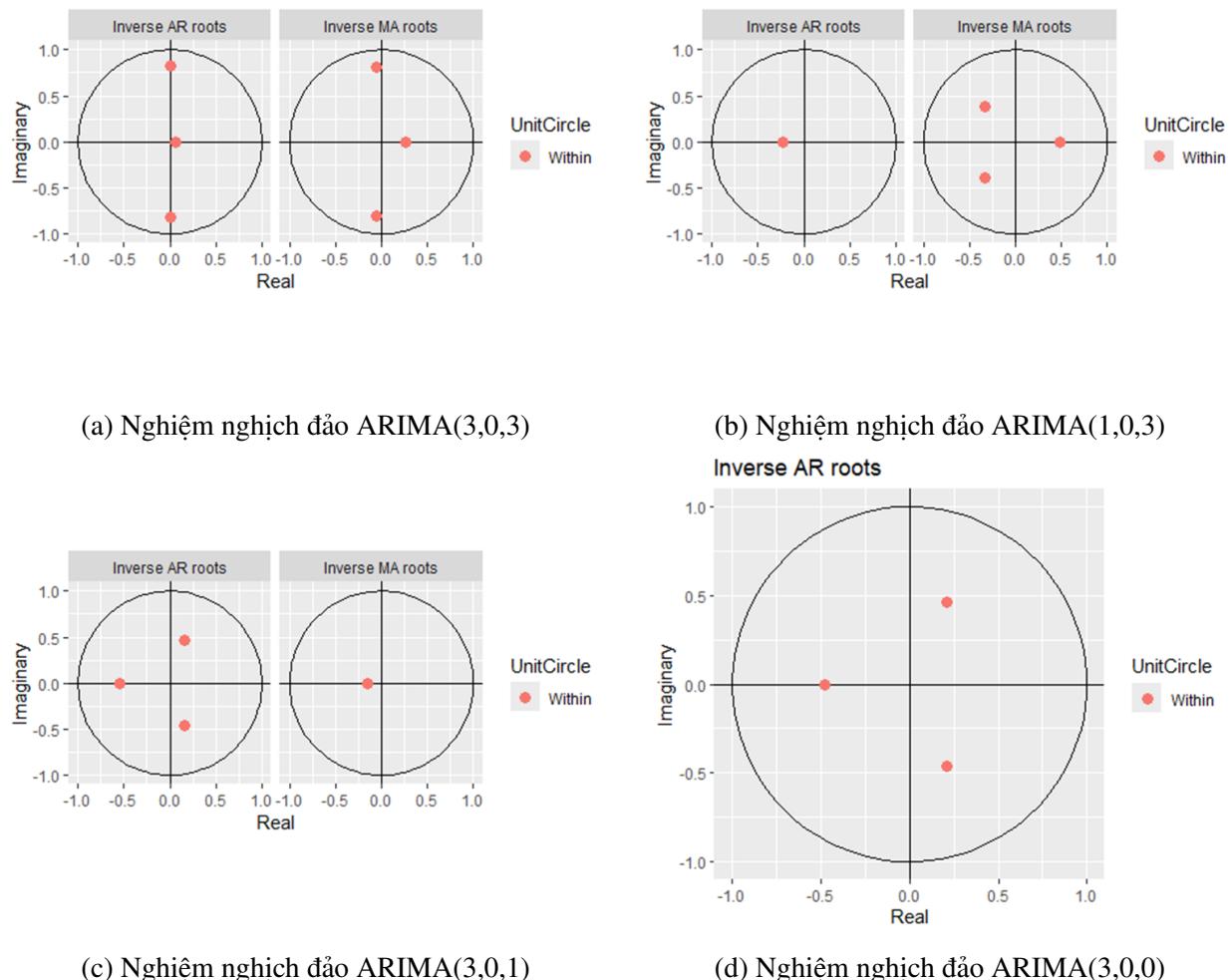
Mô hình ARIMA(3,0,0) cho chuỗi lợi suất VCB:

```

## Series: logreturn[2:499]
## ARIMA(3,0,0) with non-zero mean
##
## Coefficients:
##             ar1      ar2      ar3     mean
##            -0.0631 -0.0551 -0.1235  0.0533
## s.e.      0.0445  0.0445  0.0444  0.0429
## 
## sigma^2 = 1.425: log likelihood = -792.76
## AIC=1595.51   AICc=1595.64   BIC=1616.57
## 
## Training set error measures:
##                         ME      RMSE      MAE      MPE      MAPE      MASE
## ACF1
## Training set 0.0005231326 1.188733 0.8643201 NaN  0.6911637
## 0.002964145

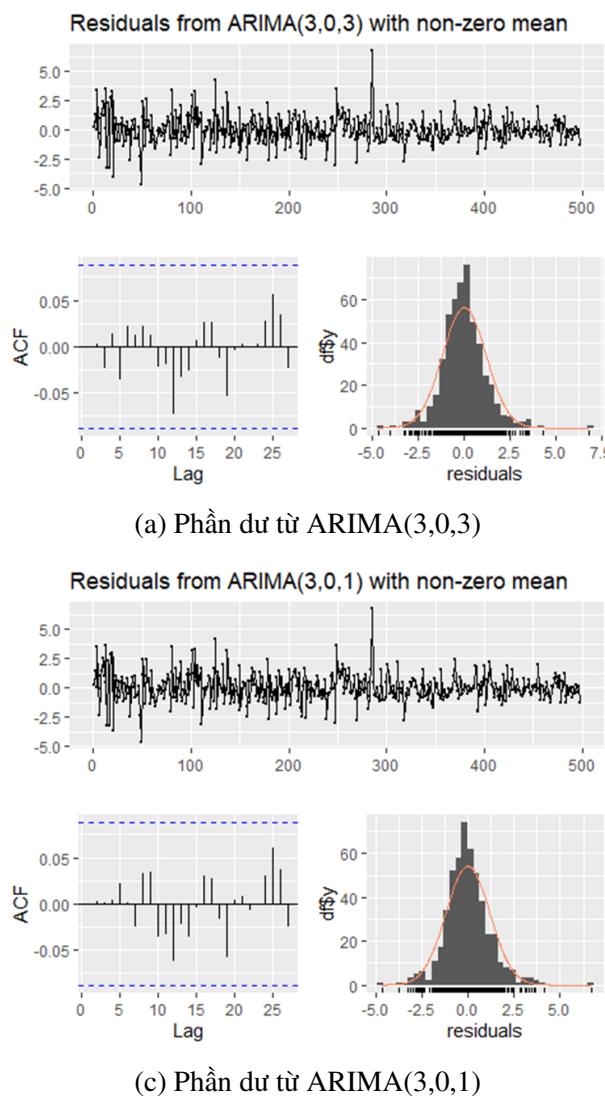
```

Bảng 5.27: Mô hình ARIMA(3,0,0) cho chuỗi lợi suất VCB

Nghiệm nghịch đảo đều nằm trong vòng tròn đơn vị:

Hình 5.3: Xem tính dừng qua nghiệm nghịch đảo

Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng:



Hình 5.4: Kiểm định tính nhiễu trắng của phân dư

Bài cá nhân - Trần Thị Kiều Oanh

Mô hình lin - lin:

```

Call:
lm(formula = DT_train ~ t_train)

Residuals:
    Min      1Q  Median      3Q     Max 
-4449.1 -2316.3 -674.4  1172.1  7616.8 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 6166.60     856.46   7.2  1.95e-09 ***  
t_train      376.54      26.14  14.4 < 2e-16 ***  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3162 on 54 degrees of freedom
Multiple R-squared:  0.7935, Adjusted R-squared:  0.7897 
F-statistic: 207.5 on 1 and 54 DF,  p-value: < 2.2e-16

```

Bảng 5.28: Kết quả hồi quy lin - lin

Mô hình lin - log:

```

Call:
lm(formula = DT_train ~ log(t_train))

Residuals:
    Min      1Q  Median      3Q     Max 
-6705.4 -3744.3 -233.1  2529.0 12115.6 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -336.8     2297.5  -0.147   0.884    
log(t_train) 5599.8      717.2   7.808 2.01e-10 ***  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4768 on 54 degrees of freedom
Multiple R-squared:  0.5303, Adjusted R-squared:  0.5216 
F-statistic: 60.97 on 1 and 54 DF,  p-value: 2.014e-10

```

Bảng 5.29: Kết quả hồi quy lin - log

Mô hình log - lin:

```

Call:
lm(formula = log(DT_train) ~ t_train)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.36664 -0.11896 -0.00374  0.07454  0.42426 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 9.028276   0.047356 190.65 <2e-16 ***  
t_train     0.022050   0.001445   15.26 <2e-16 ***  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1748 on 54 degrees of freedom
Multiple R-squared:  0.8117, Adjusted R-squared:  0.8082 
F-statistic: 232.7 on 1 and 54 DF,  p-value: < 2.2e-16

```

Bảng 5.30: Kết quả hồi quy log - lin

Mô hình log - log:

```

Call:
lm(formula = log(DT_train) ~ log(t_train))

Residuals:
    Min      1Q  Median      3Q     Max 
-0.40942 -0.17466  0.03677  0.12649  0.45567 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 8.55845   0.11608  73.729 < 2e-16 ***  
log(t_train) 0.35684   0.03624   9.848 1.17e-13 ***  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2409 on 54 degrees of freedom
Multiple R-squared:  0.6423, Adjusted R-squared:  0.6357 
F-statistic: 96.97 on 1 and 54 DF,  p-value: 1.175e-13

```

Bảng 5.31: Kết quả hồi quy log - log

Mô hình hồi quy với biến giả mùa vụ:

```

Call:
lm(formula = DT_train ~ s2 + s3 + s4)

Residuals:
    Min      1Q  Median      3Q     Max 
-10270  -5734  -1564   3779  17264 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 16175.7     1890.2   8.558 1.69e-11 ***  
s2          676.9      2673.1   0.253   0.801    
s3          858.7      2673.1   0.321   0.749    
s4         1353.6      2673.1   0.506   0.615    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 7072 on 52 degrees of freedom
Multiple R-squared:  0.005039, Adjusted R-squared:  -0.05236 
F-statistic: 0.08779 on 3 and 52 DF,  p-value: 0.9664

```

Bảng 5.32: Kết quả hồi quy với biến giả mùa vụ

Xu thế tuyến tính + Mùa vụ, dạng cộng:

```

Call:
lm(formula = DT_train ~ t_train + s2 + s3 + s4)

Residuals:
    Min      1Q  Median      3Q     Max 
-4517.0 -2310.2 -687.2  1175.5  7479.9 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 6015.32     1133.25   5.308 2.43e-06 ***  
t_train      376.31      26.94   13.966 < 2e-16 ***  
s2          300.57     1229.13   0.245   0.808    
s3          106.03     1230.02   0.086   0.932    
s4          224.68     1231.49   0.182   0.856    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 3251 on 51 degrees of freedom
Multiple R-squared:  0.7938, Adjusted R-squared:  0.7776 
F-statistic: 49.08 on 4 and 51 DF,  p-value: < 2.2e-16

```

Bảng 5.33: Kết quả xu thế tuyến tính có mùa vụ, dạng cộng

Xu thế tuyến tính + Mùa vụ , dạng nhân:

```

Call:
lm(formula = DT_train ~ t_train + I(t_train * s2) + I(t_train *
s3) + I(t_train * s4))

Residuals:
    Min      1Q  Median      3Q     Max 
-4565.8 -2206.3 -681.5 1135.4 7419.3 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 6174.995   880.663   7.012 5.21e-09 ***
t_train      368.514    36.612  10.065 1.03e-13 ***
I(t_train * s2) 11.527    38.534   0.299   0.766  
I(t_train * s3)  6.296    38.052   0.165   0.869  
I(t_train * s4) 12.536    37.604   0.333   0.740  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3249 on 51 degrees of freedom
Multiple R-squared:  0.7941, Adjusted R-squared:  0.7779 
F-statistic: 49.16 on 4 and 51 DF,  p-value: < 2.2e-16

```

Bảng 5.34: Kết quả xu thế tuyến tính có mùa vụ dạng nhân

Xu thế không tuyến tính + Mùa vụ , dạng cộng :

```

Call:
lm(formula = log(DT_train) ~ t_train + s2 + s3 + s4)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.35180 -0.12553 -0.01323  0.07221  0.42302 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 9.013454   0.062607 143.970  <2e-16 ***
t_train      0.022024   0.001489  14.796  <2e-16 ***
s2          0.025128   0.067904   0.370   0.713  
s3          0.016236   0.067953   0.239   0.812  
s4          0.020847   0.068034   0.306   0.761  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1796 on 51 degrees of freedom
Multiple R-squared:  0.8123, Adjusted R-squared:  0.7975 
F-statistic: 55.16 on 4 and 51 DF,  p-value: < 2.2e-16

```

Bảng 5.35: Kết quả xu thế không tuyến tính có mùa vụ dạng cộng

Xu thế không tuyến tính + Mùa vụ , dạng nhân :

```

Call:
lm(formula = log(DT_train) ~ t_train + I(t_train * s2) + I(t_train *
s3) + I(t_train * s4))

Residuals:
    Min      1Q  Median      3Q     Max 
-0.36652 -0.12061 -0.01139  0.07504  0.42467 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 9.0283424  0.0487387 185.240 < 2e-16 ***
t_train      0.0218593  0.0020262  10.788 9.11e-15 ***
I(t_train * s2) 0.0004276  0.0021326   0.200   0.842  
I(t_train * s3) 0.0001216  0.0021059   0.058   0.954  
I(t_train * s4) 0.0001985  0.0020811   0.095   0.924  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1798 on 51 degrees of freedom
Multiple R-squared:  0.8118, Adjusted R-squared:  0.7971 
F-statistic: 55.01 on 4 and 51 DF,  p-value: < 2.2e-16

```

Bảng 5.36: Kết quả xu thế không tuyến tính có mùa vụ dạng nhân

Holt-Winter có mùa vụ dạng cộng :

```

Holt-Winters exponential smoothing with trend and additive seasonal component.

Call:
HoltWinters(x = DT_train, seasonal = "a")

Smoothing parameters:
alpha: 1
beta : 0.4931384
gamma: 0.9361004

Coefficients:
[,1]
a 33027.0349
b 261.3418
s1 192.8182
s2 446.6110
s3 -249.6323
s4 -389.7969

```

Bảng 5.37: Kết quả Holt-Winter có mùa vụ dạng cộng

Holt-Winter có mùa vụ dạng nhân:

```
Holt-Winters exponential smoothing with trend and multiplicative seasonal component.

Call:
HoltWinters(x = DT_train, seasonal = "m")

Smoothing parameters:
alpha: 0.9682685
beta : 0.3398843
gamma: 1

Coefficients:
[,1]
a 3.371011e+04
b 8.651010e+02
s1 1.005512e+00
s2 1.032101e+00
s3 9.898563e-01
s4 9.681736e-01
```

Bảng 5.38: Kết quả Holt-Winter có mùa vụ dạng nhân

Kết quả các mô hình ARIMA cho chuỗi giá đóng cửa CTG ARIMA(1,1,1)

```
Series: price
ARIMA(1,1,1) with drift

Coefficients:
ar1     ma1     drift
-0.4373  0.2821  25.7488
s.e.    0.1511  0.1579  19.7918

sigma^2 = 246566: log likelihood = -3796.57
AIC=7601.15   AICC=7601.23   BIC=7617.99

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.1257538 494.5601 345.4791 -0.01811645 1.139784 0.9987497 0.01198832
```

ARIMA(4,1,4)

ARIMA(4,1,4) with drift

```
Coefficients:
ar1     ar2     ar3     ar4     ma1     ma2     ma3     ma4     drift
-0.1089  0.2357  0.303   -0.1958  -0.0276  -0.1328  -0.3163  0.1018  25.3971
s.e.    0.5227  0.3143  0.286   0.3014   0.5254   0.2891   0.2716  0.3045  17.9130

sigma^2 = 243312: log likelihood = -3790.28
AIC=7600.57   AICC=7601.02   BIC=7642.68

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.4094675 488.2993 345.7504 -0.01813961 1.140686 0.9995337 -0.0004927134
```

ARIMA(1,1,4)

ARIMA(1,1,4) with drift

Coefficients:

| | ar1 | ma1 | ma2 | ma3 | ma4 | drift |
|------|---------|--------|--------|--------|---------|---------|
| s.e. | -0.5325 | 0.3926 | 0.0428 | 0.0078 | -0.1282 | 25.6306 |
| | 0.2943 | 0.2939 | 0.0597 | 0.0592 | 0.0432 | 18.8464 |

$\sigma^2 = 242828$: log likelihood = -3791.29
AIC=7596.59 AICc=7596.81 BIC=7626.06

Training set error measures:

| | ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
|--------------|----------|----------|----------|------------|----------|-----------|--------------|
| Training set | 0.305152 | 489.3074 | 345.8863 | -0.0179134 | 1.141311 | 0.9999267 | 0.0006555241 |

ARIMA(4,1,1)

ARIMA(4,1,1) with drift

Coefficients:

| | ar1 | ar2 | ar3 | ar4 | ma1 | drift |
|------|---------|--------|---------|---------|--------|---------|
| s.e. | -0.3368 | 0.0746 | -0.0020 | -0.1380 | 0.1981 | 25.6190 |
| | 0.2785 | 0.0616 | 0.0537 | 0.0452 | 0.2797 | 18.7501 |

$\sigma^2 = 242399$: log likelihood = -3790.86
AIC=7595.72 AICc=7595.95 BIC=7625.19

Training set error measures:

| | ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
|--------------|-----------|----------|----------|-------------|----------|----------|--------------|
| Training set | 0.3123405 | 488.8751 | 346.6869 | -0.01792226 | 1.143833 | 1.002241 | 0.0005228244 |

Kết quả các mô hình ARIMA cho chuỗi log return giá CTG

ARIMA(0,0,4)

Series: Data1\$`Log-return(%)`
ARIMA(0,0,4) with non-zero mean

Coefficients:

| | ma1 | ma2 | ma3 | ma4 | mean |
|------|---------|--------|---------|---------|--------|
| s.e. | -0.1331 | 0.1161 | -0.0554 | -0.0848 | 0.0748 |
| | 0.0444 | 0.0451 | 0.0470 | 0.0451 | 0.0592 |

$\sigma^2 = 2.526$: log likelihood = -953.74
AIC=1919.49 AICc=1919.65 BIC=1944.87

Training set error measures:

| | ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
|--------------|--------------|----------|----------|-----|------|-----------|--------------|
| Training set | 0.0009138504 | 1.581583 | 1.138357 | NaN | Inf | 0.6507336 | -0.001300926 |

ARIMA(2,0,2)

Series: Data1\$`Log-return(%)`
ARIMA(2,0,2) with non-zero mean

Coefficients:

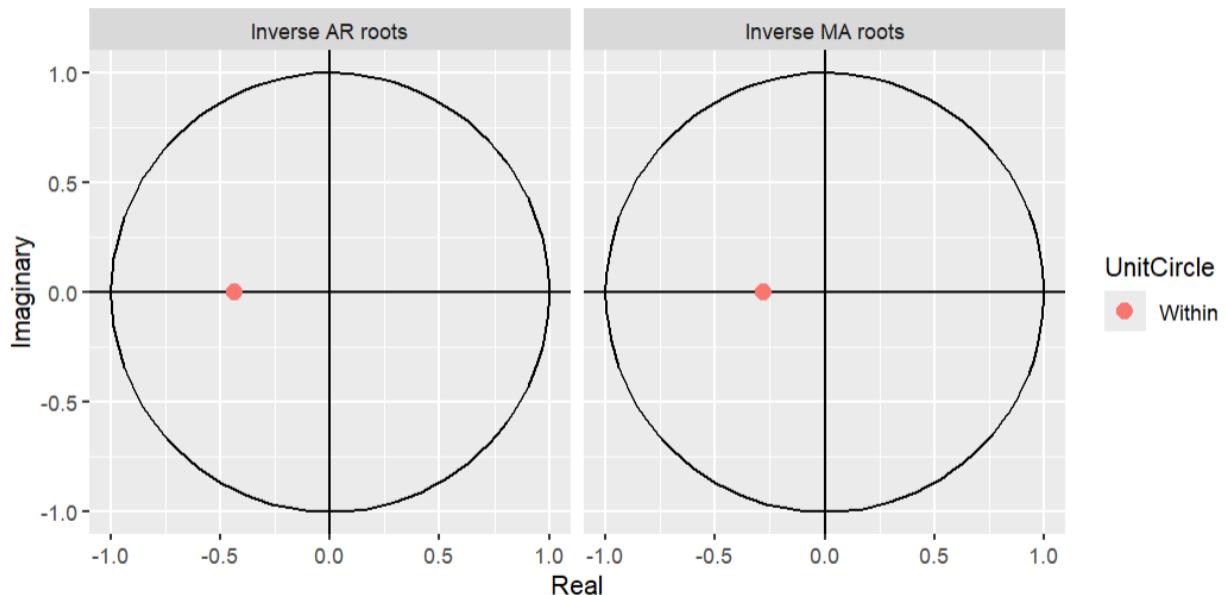
| | ar1 | ar2 | ma1 | ma2 | mean |
|------|---------|---------|--------|--------|--------|
| | -0.9717 | -0.6360 | 0.8397 | 0.6326 | 0.0742 |
| s.e. | 0.2236 | 0.1519 | 0.2173 | 0.1351 | 0.0665 |

$\sigma^2 = 2.526$: log likelihood = -953.73
AIC=1919.46 AICc=1919.63 BIC=1944.84

Training set error measures:

| ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
|---------------------------|----------|----------|-----|------|----------|-------------|
| Training set 0.0007397522 | 1.581522 | 1.132645 | NaN | Inf | 0.647468 | 0.003649015 |

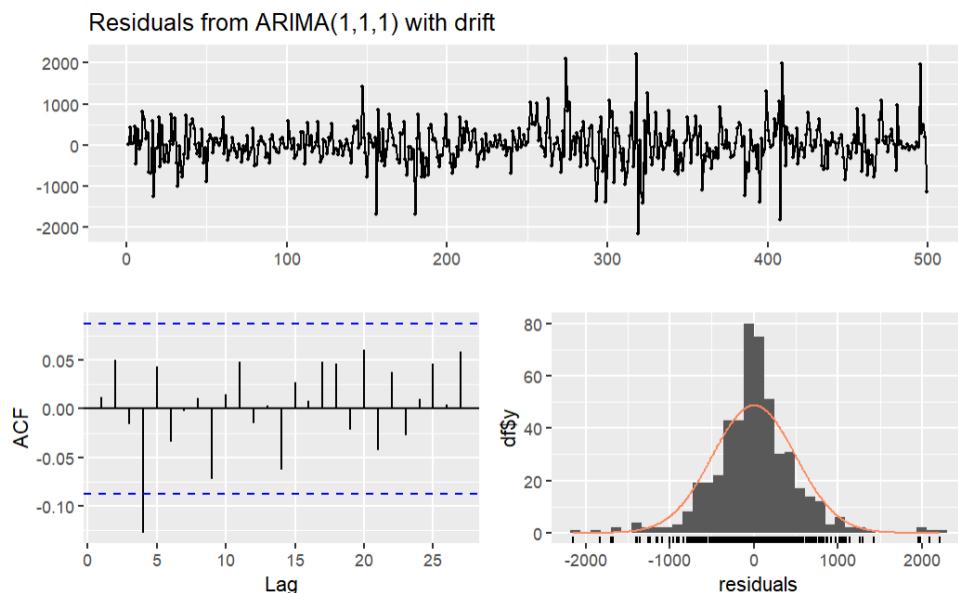
Kiểm định về mô hình ARIMA(1,1,1)



Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị

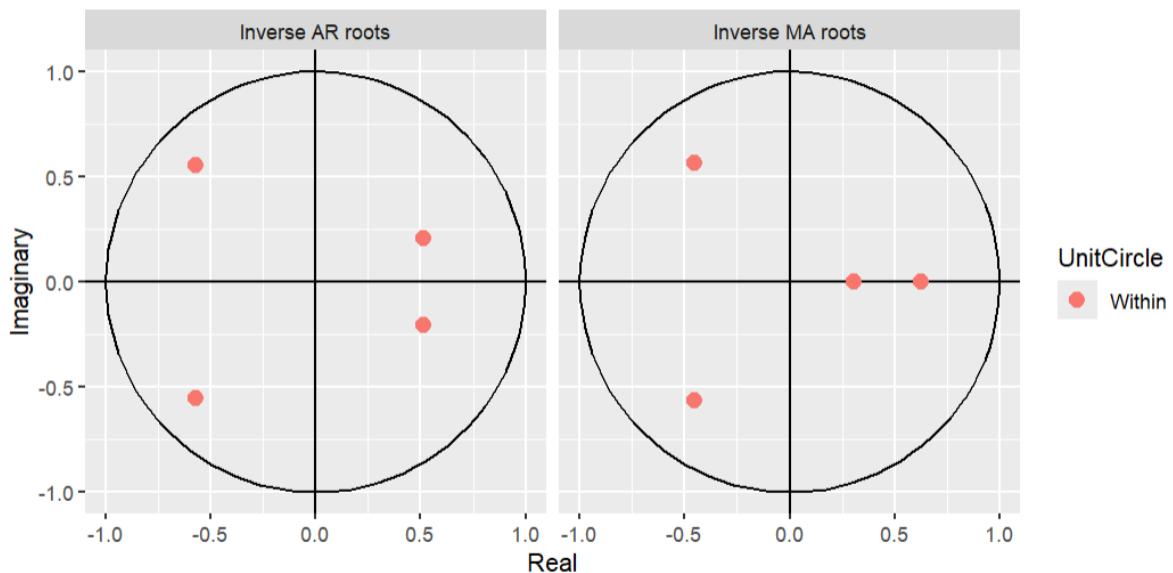
Ljung-Box test

```
data: Residuals from ARIMA(1,1,1) with drift
Q* = 14.053, df = 8, p-value = 0.08038
Model df: 2. Total lags used: 10
```



$p - Value = 0.08038 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

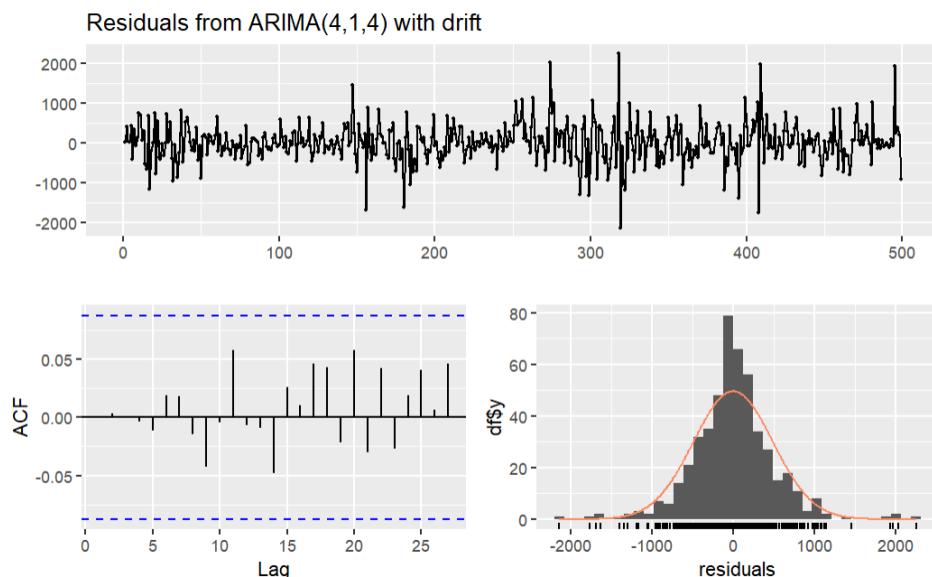
ARIMA(4,1,4)



Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị

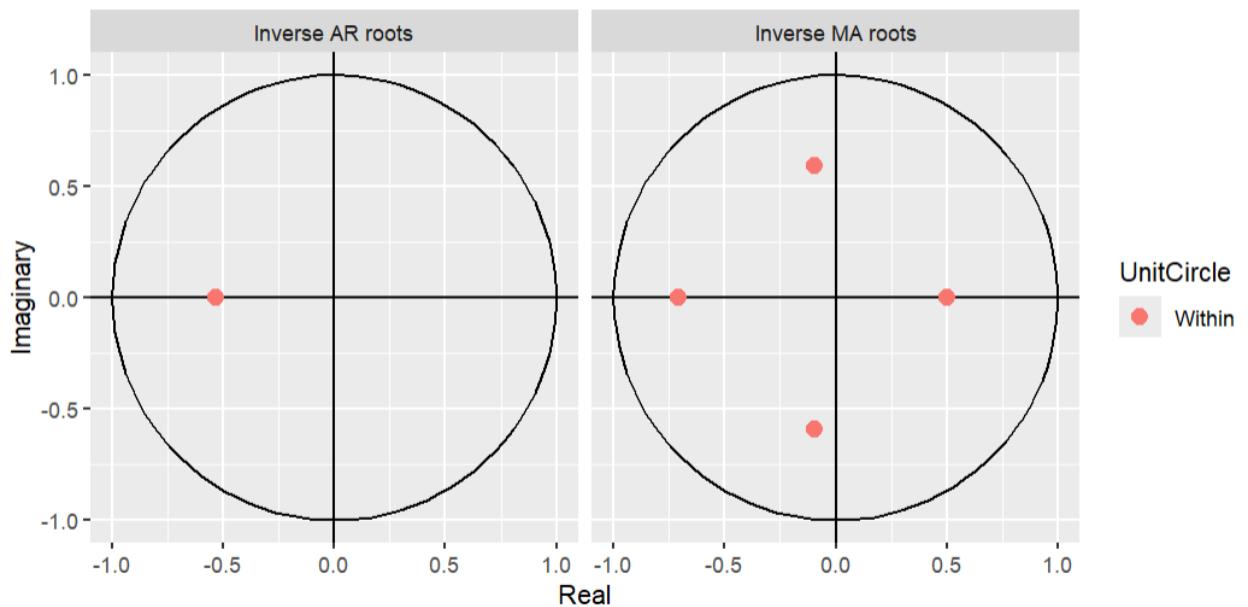
```
Ljung-Box test
data: Residuals from ARIMA(4,1,4) with drift
Q* = 3.1894, df = 3, p-value = 0.3633

Model df: 8.  Total lags used: 11
```



$p - Value = 0.3633 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

ARIMA(1,1,4)

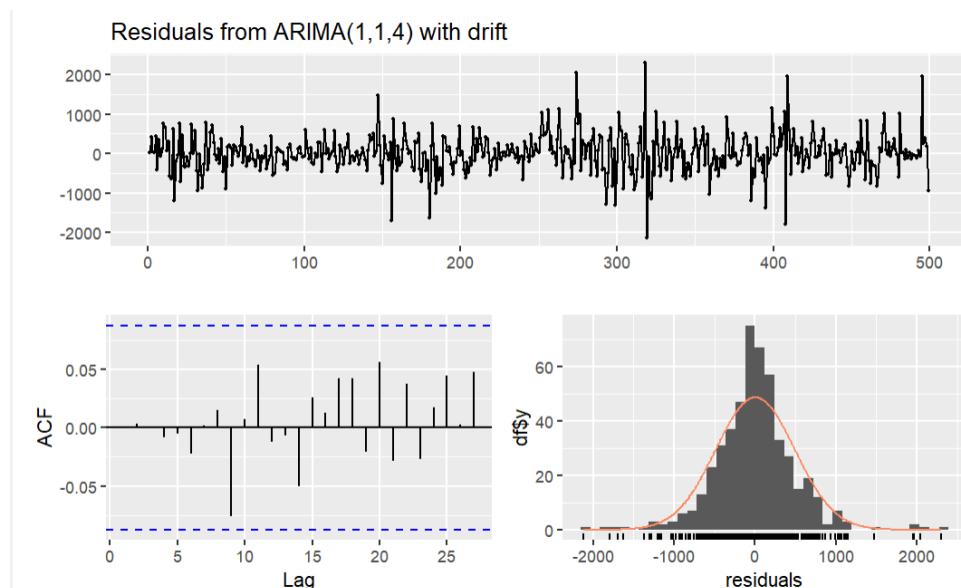


Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị

Ljung-Box test

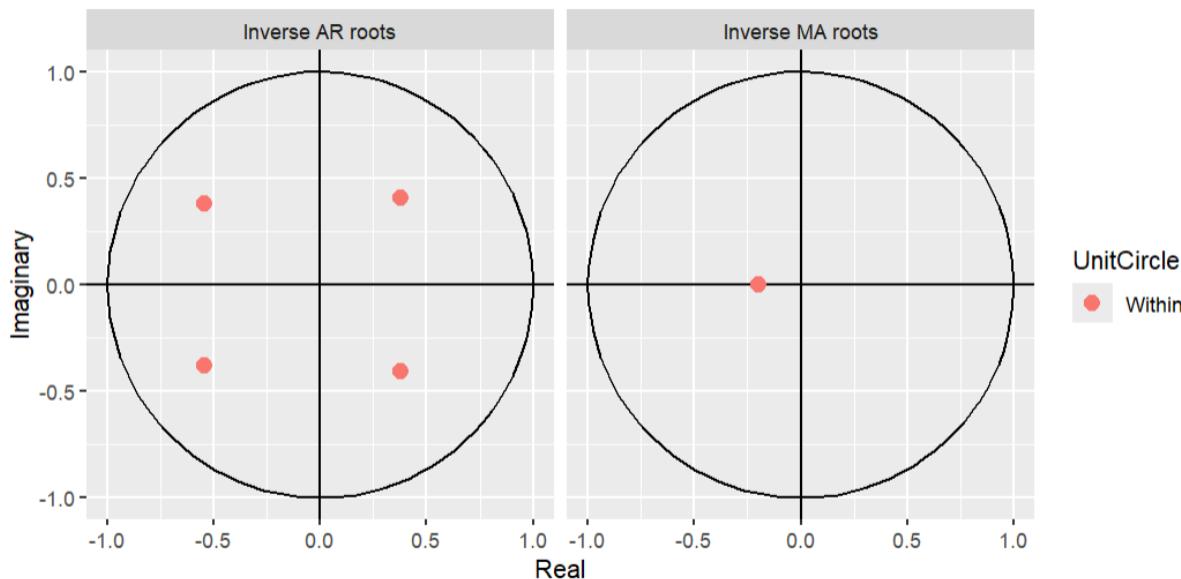
```
data: Residuals from ARIMA(1,1,4) with drift
Q* = 3.4321, df = 5, p-value = 0.6337
```

```
Model df: 5. Total lags used: 10
```



$p - Value = 0.6337 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

ARIMA(4,1,1)

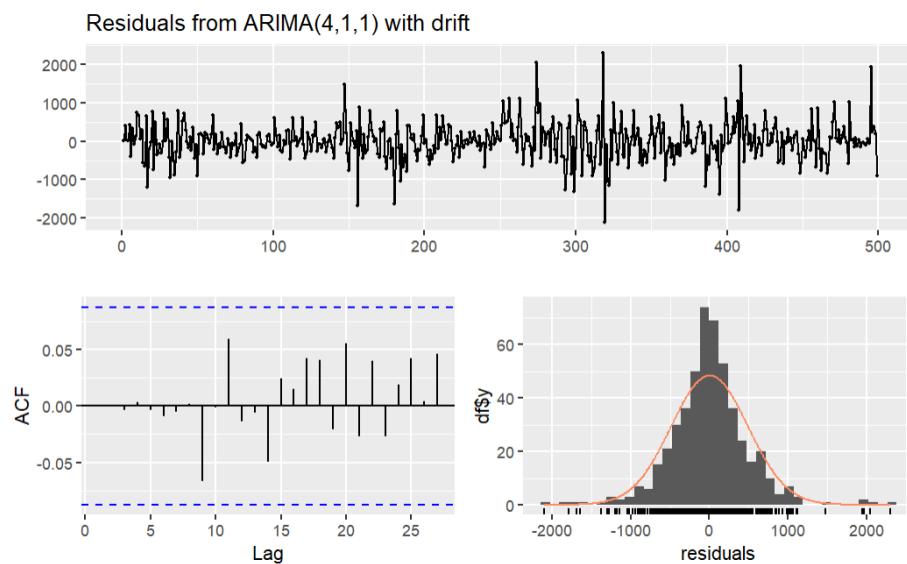


Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị

```
Ljung-Box test

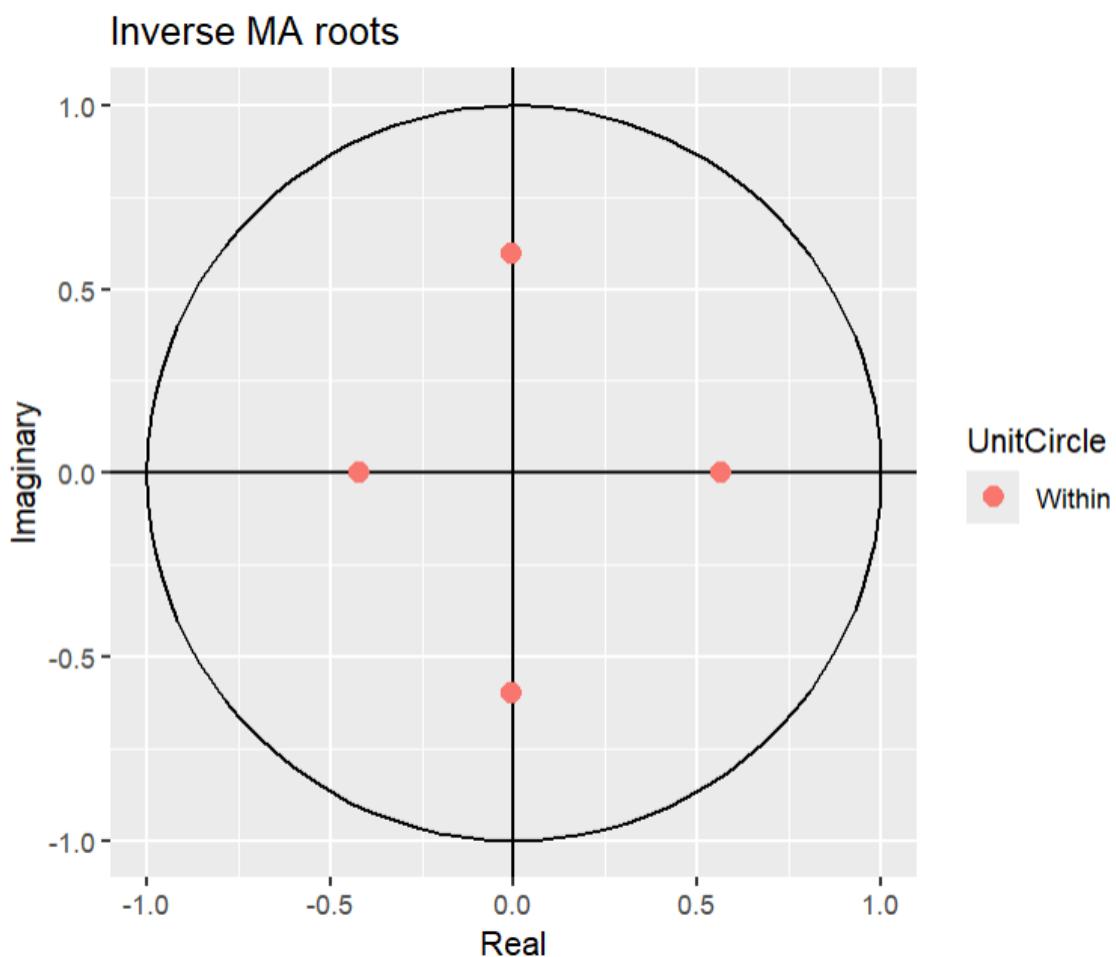
data: Residuals from ARIMA(4,1,1) with drift
Q* = 2.3508, df = 5, p-value = 0.7988

Model df: 5.  Total lags used: 10
```



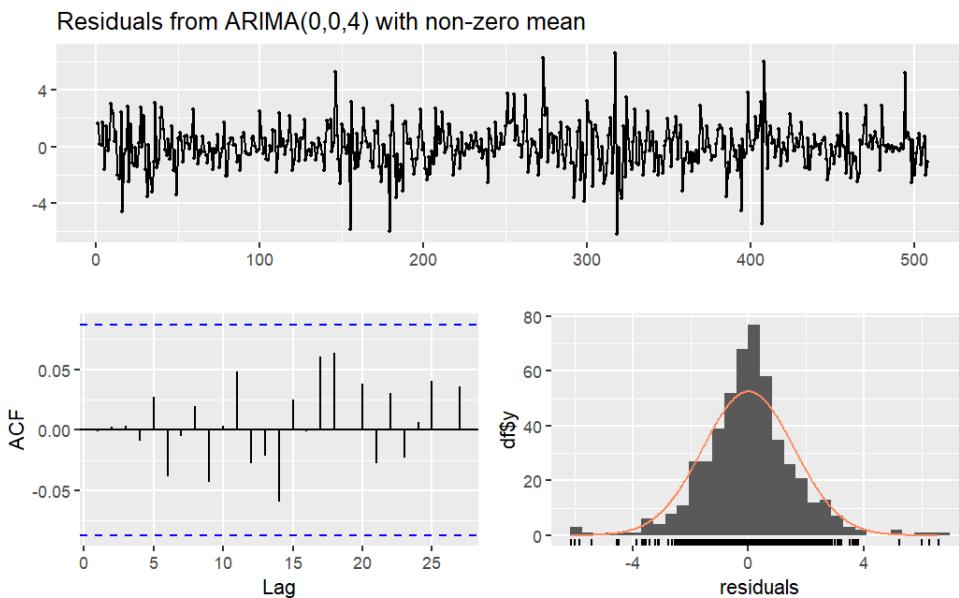
$p - Value = 0.7988 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

ARIMA(0,0,4)



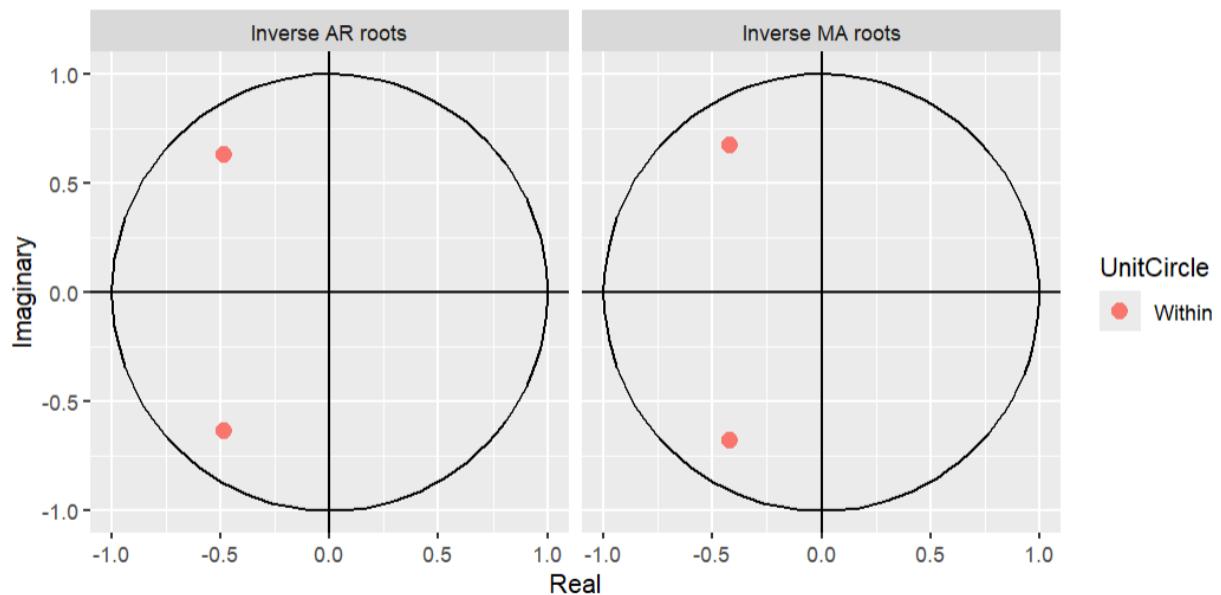
Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị

```
Ljung-Box test
data: Residuals from ARIMA(0,0,4) with non-zero mean
Q* = 2.3683, df = 6, p-value = 0.8829
Model df: 4. Total lags used: 10
```



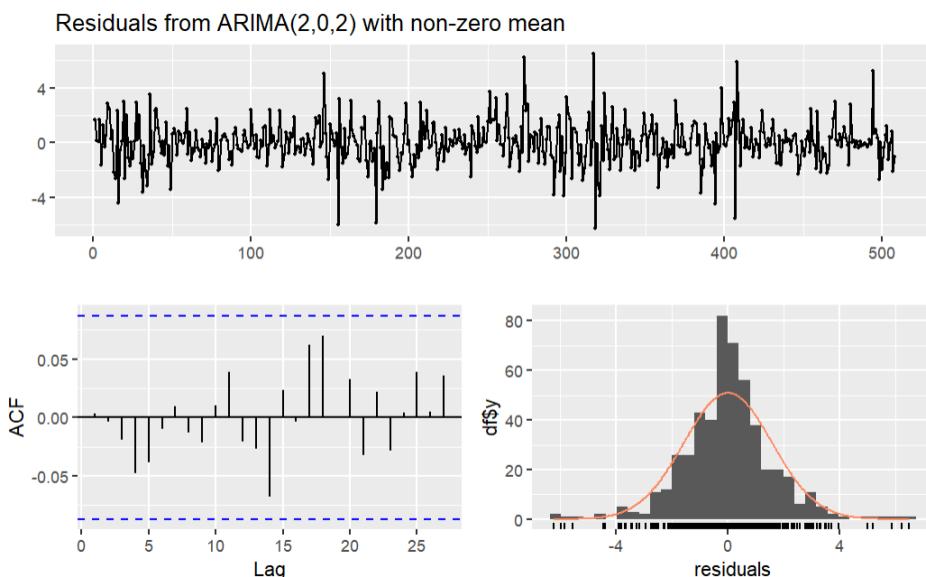
$p - Value = 0.8829 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

ARIMA(2,0,2)



Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị

```
Ljung-Box test
data: Residuals from ARIMA(2,0,2) with non-zero mean
Q* = 2.6144, df = 6, p-value = 0.8555
Model df: 4. Total lags used: 10
```



$p - Value = 0.8555 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

Bài cá nhân - Nghiêm Gia Phương

CHUỖI DOANH THU

Mô hình lin-lin

```

Call:
lm(formula = dt_train ~ t_train)

Residuals:
    Min      1Q  Median      3Q     Max 
-584.06 -173.42   -4.79  171.72  602.88 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 792.395    67.953 11.661 2.25e-16 ***
t_train     -1.278     2.074 -0.616    0.54    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 250.9 on 54 degrees of freedom
Multiple R-squared:  0.006981, Adjusted R-squared: -0.01141 
F-statistic: 0.3796 on 1 and 54 DF,  p-value: 0.5404

```

Hình 5.5: Kết quả mô hình lin-lin

Mô hình lin-log

```

Call:
lm(formula = dt_train ~ log(t_train))

Residuals:
    Min      1Q  Median      3Q     Max 
-481.08 -182.31 -12.28 179.23 623.33 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 640.44     120.18   5.329 1.98e-06 ***
log(t_train) 37.54      37.52   1.001    0.321  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 249.4 on 54 degrees of freedom
Multiple R-squared:  0.0182, Adjusted R-squared:  2.174e-05 
F-statistic: 1.001 on 1 and 54 DF, p-value: 0.3215

```

Mô hình log-lin

```

Call:
lm(formula = log(dt_train) ~ t_train)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.23938 -0.21458  0.05566  0.27752  0.65645 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 6.572638   0.103791  63.326 <2e-16 ***
t_train     -0.000270   0.003168  -0.085    0.932  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3832 on 54 degrees of freedom
Multiple R-squared:  0.0001345, Adjusted R-squared:  -0.01838 
F-statistic: 0.007262 on 1 and 54 DF, p-value: 0.9324

```

Mô hình log-log

```

Call:
lm(formula = log(dt_train) ~ log(t_train))

Residuals:
    Min      1Q  Median      3Q     Max 
-0.9267 -0.2260  0.0374  0.2944  0.6781 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 6.2567     0.1794  34.879 <2e-16 ***
log(t_train) 0.1002     0.0560   1.789   0.0793 . 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3723 on 54 degrees of freedom
Multiple R-squared:  0.05593, Adjusted R-squared:  0.03845 
F-statistic: 3.199 on 1 and 54 DF, p-value: 0.07927

```

Mô hình hồi quy với biến giả mùa vụ

```

Call:
lm(formula = dt_train ~ s2 + s3 + s4)

Residuals:
    Min      1Q  Median      3Q     Max 
-493.84 -143.96 -12.19  180.45  600.78 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 584.28     61.77   9.459 6.83e-13 ***
s2          187.21     87.35   2.143  0.03680 *  
s3          204.33     87.35   2.339  0.02321 *  
s4          295.24     87.35   3.380  0.00138 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 231.1 on 52 degrees of freedom
Multiple R-squared:  0.1884, Adjusted R-squared:  0.1416 
F-statistic: 4.023 on 3 and 52 DF,  p-value: 0.01197

```

Mô hình xu thế tuyến tính+mùa vụ dạng cộng

```

Call:
lm(formula = dt_train ~ t_train + s2 + s3 + s4)

Residuals:
    Min      1Q  Median      3Q     Max 
-538.50 -139.96 -16.56  163.50  583.60 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 630.669     80.711   7.814 2.84e-10 ***
t_train     -1.718      1.919  -0.895  0.37484  
s2          188.930     87.540   2.158  0.03564 *  
s3          207.764     87.603   2.372  0.02152 *  
s4          300.393     87.708   3.425  0.00122 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 231.6 on 51 degrees of freedom
Multiple R-squared:  0.2009, Adjusted R-squared:  0.1383 
F-statistic: 3.206 on 4 and 51 DF,  p-value: 0.02008

```

Mô hình xu thế tuyến tính+mùa vụ dạng nhân

```

Call:
lm(formula = dt_train ~ t_train + I(t_train * s2) + I(t_train * 
s3) + I(t_train * s4))

Residuals:
    Min      1Q  Median      3Q     Max 
-585.47 -141.98   13.16  172.55  593.06 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 799.186     63.872 12.512 < 2e-16 ***
t_train     -6.658      2.655  -2.508  0.01538 *  
I(t_train * s2) 5.549      2.795   1.985  0.05249 .  
I(t_train * s3) 6.033      2.760   2.186  0.03342 *  
I(t_train * s4) 8.530      2.727   3.128  0.00291 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 235.6 on 51 degrees of freedom
Multiple R-squared:  0.1725, Adjusted R-squared:  0.1076 
F-statistic: 2.657 on 4 and 51 DF,  p-value: 0.04323

```

Mô hình Holt - Winter có mùa vụ dạng cộng

```
Holt-Winters exponential smoothing with trend and additive seasonal component.

Call:
HoltWinters(x = dt_train, seasonal = "a")

Smoothing parameters:
alpha: 0.3250884
beta : 0.165585
gamma: 0.7236907

Coefficients:
[,1]
a 934.19825
b 11.71475
s1 -296.51479
s2 -158.82354
s3 -105.98938
s4 86.79820
```

Mô hình Holt - Winter có mùa vụ dạng nhân

```
Holt-Winters exponential smoothing with trend and multiplicative seasonal component.

Call:
HoltWinters(x = dt_train, seasonal = "m")

Smoothing parameters:
alpha: 0.3600538
beta : 0.1372853
gamma: 1

Coefficients:
[,1]
a 1152.5983021
b 17.6979748
s1 0.5630960
s2 0.6406033
s3 0.7181630
s4 0.8772290
```

Mô hình xu thế tuyến tính + Mùa vụ, dạng nhân trên toàn dữ liệu 2010-2024

```
Call:
lm(formula = doanhthu ~ time + I(time * s2) + I(time * s3) +
   I(time * s4))

Residuals:
    Min      1Q  Median      3Q     Max 
-581.33 -141.65  15.89  127.28  591.76 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  795.007    59.496 13.362 < 2e-16 ***
time        -6.625     2.306 -2.873 0.005770 **  
I(time * s2)  5.820     2.430  2.395 0.020040 *   
I(time * s3)  6.063     2.401  2.525 0.014482 *   
I(time * s4)  8.923     2.374  3.758 0.000416 *** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 227.4 on 55 degrees of freedom
Multiple R-squared:  0.2144,    Adjusted R-squared:  0.1573 
F-statistic: 3.753 on 4 and 55 DF,  p-value: 0.009046
```

CHUỖI GIÁ VÀ LOG-RETURN CỦA PVC

**Kết quả các mô hình ARIMA
ARIMA(1,0,5)**

```

Call:
arima(x = logreturn[2:499], order = c(1, 0, 5), include.mean = FALSE)

Coefficients:
      ar1     ma1     ma2     ma3     ma4     ma5 
    -0.6044  0.5231 -0.1516 -0.0110 -0.0144  0.1297
  s.e.  0.1474  0.1495  0.0512  0.0565  0.0520  0.0559 

sigma^2 estimated as 6.035:  log likelihood = -1154.38,  aic = 2322.76

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1 
Training set 0.01674591 2.456576 1.696338 NaN   Inf  0.6224458 -0.006078244

```

ARIMA(5,0,1)

```

Call:
arima(x = logreturn[2:499], order = c(5, 0, 1), include.mean = FALSE)

Coefficients:
      ar1     ar2     ar3     ar4     ar5     ma1 
    -0.7550 -0.1632 -0.0349 -0.0366  0.1159  0.6701
  s.e.  0.1187  0.0583  0.0581  0.0567  0.0521  0.1139 

sigma^2 estimated as 6.007:  log likelihood = -1153.21,  aic = 2320.41

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1 
Training set 0.01749898 2.450881 1.694231 NaN   Inf  0.6216724 0.001030843

```

ARIMA(1,1,5)

```

Call:
arima(x = price, order = c(1, 1, 5))

Coefficients:
      ar1     ma1     ma2     ma3     ma4     ma5 
    -0.6261  0.5545 -0.1566  0.0000 -0.0036  0.1396
  s.e.  0.1222  0.1244  0.0512  0.0568  0.0528  0.0548 

sigma^2 estimated as 103664:  log likelihood = -3582.54,  aic = 7179.07

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1 
Training set 1.643253 321.6463 223.7773 -0.01551058 1.703251 0.9917035 -0.007314143

```

ARIMA(5,1,1)

```

Call:
arima(x = price, order = c(5, 1, 1))

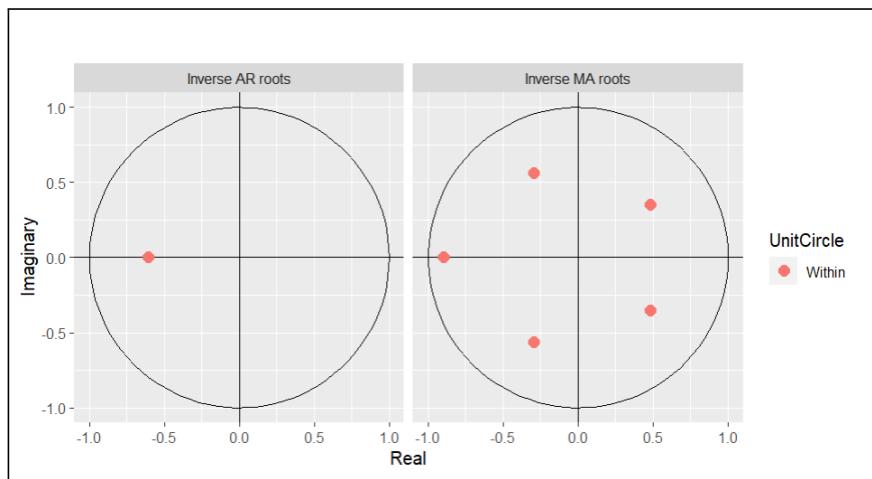
Coefficients:
      ar1     ar2     ar3     ar4     ar5     ma1 
    -0.7880 -0.1709 -0.0290 -0.0218  0.1303  0.7102
  s.e.  0.1048  0.0584  0.0584  0.0572  0.0516  0.0992 

sigma^2 estimated as 103173:  log likelihood = -3581.31,  aic = 7176.62

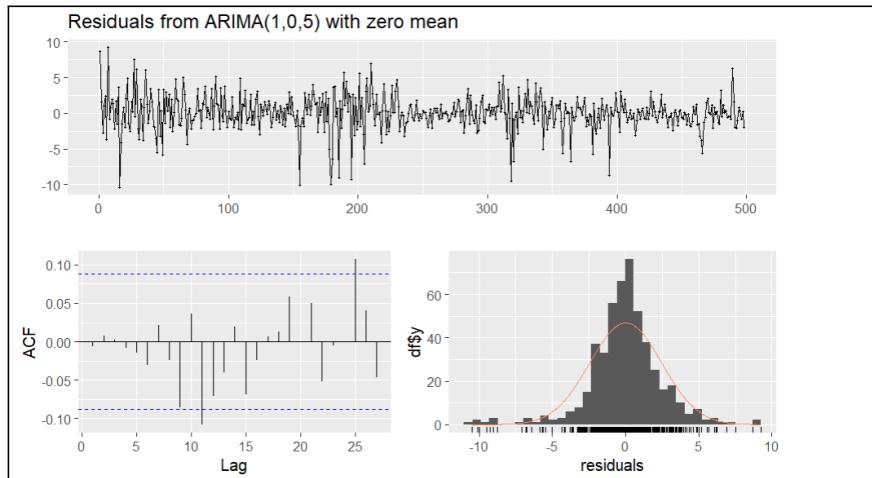
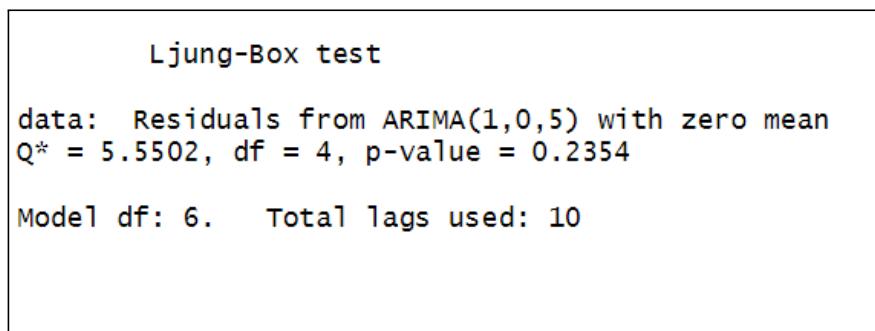
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1 
Training set 1.725283 320.8836 223.3879 -0.01550499 1.699545 0.9899777 0.001331992

```

**Kết quả kiểm định các mô hình
ARIMA(1,0,5)**

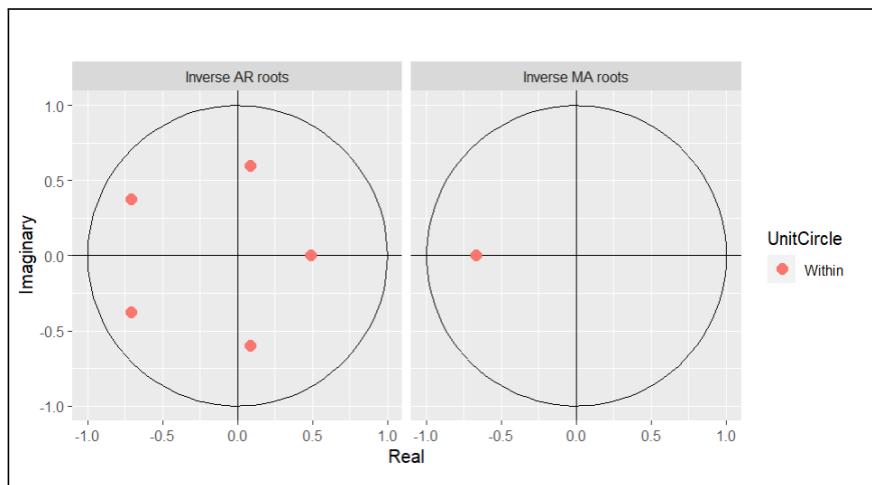


Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị.

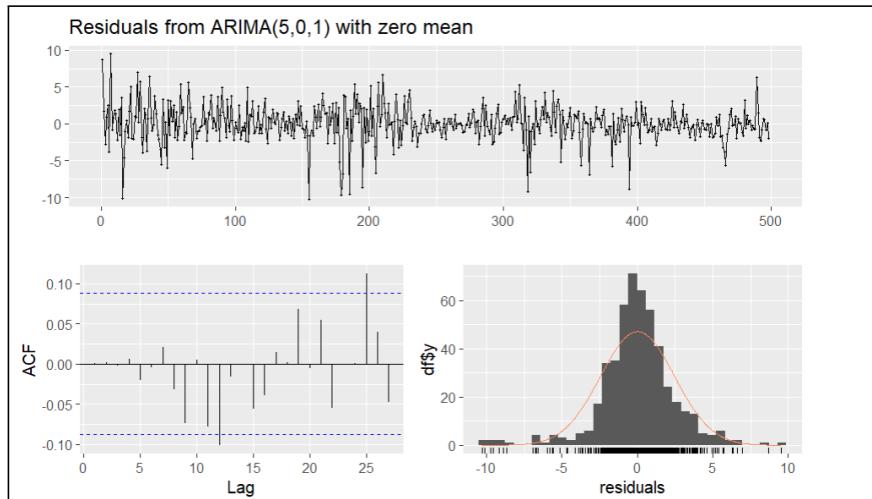
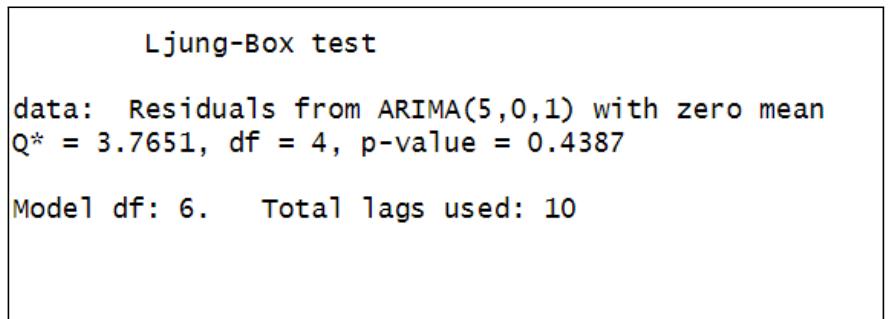


$p-value = 0.2354 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

ARIMA(5,0,1)

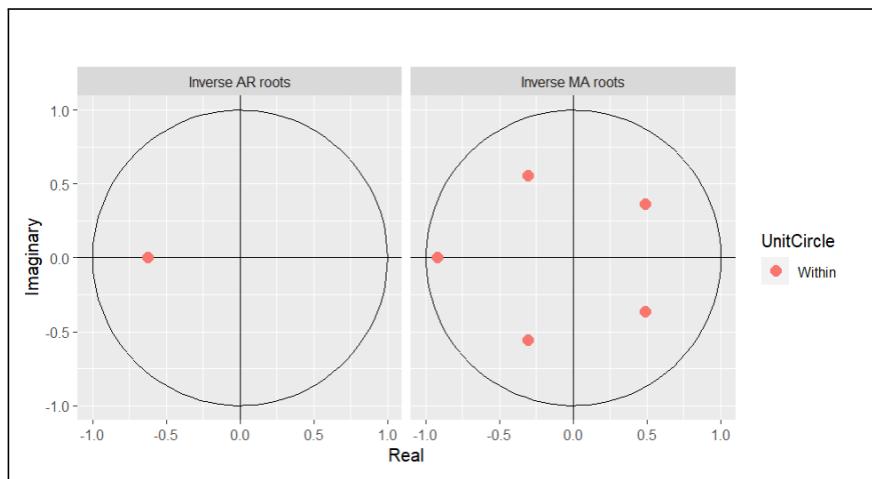


Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị.

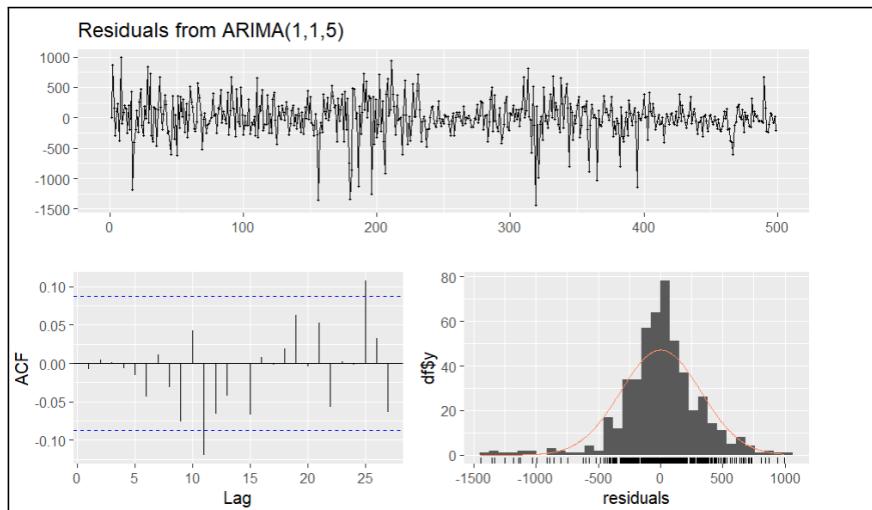
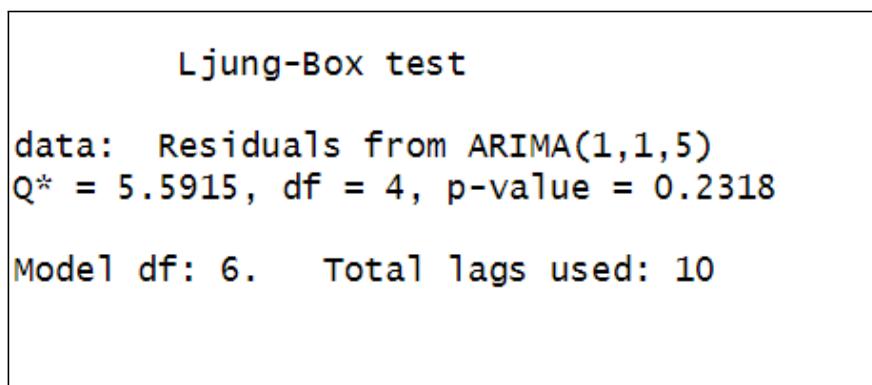


$p - value = 0.4387 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

ARIMA(1,1,5)

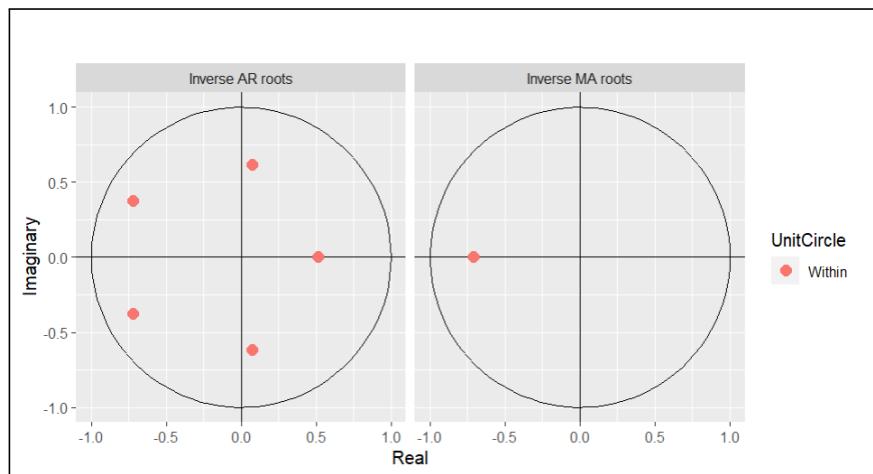


Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị.

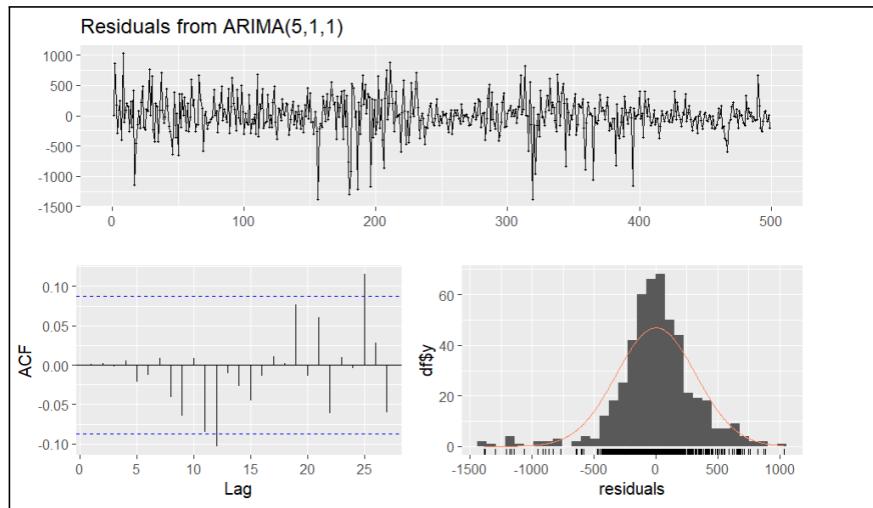
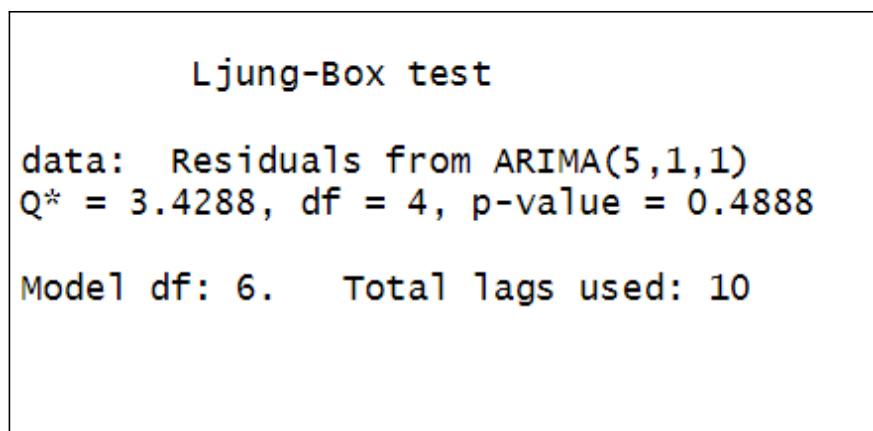


$p - value = 0.2318 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

ARIMA(5,1,1)



Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị.



$p-value = 0.4888 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

MÔ HÌNH GARCH

* ----- *

```

*          GARCH Model Fit          *
*-----*

Conditional Variance Dynamics
-----
GARCH Model : sGARCH(1,1)
Mean Model : ARFIMA(5,0,1)
Distribution : norm

Optimal Parameters
-----
            Estimate Std. Error   t value Pr(>|t|)
mu      -0.060731  0.081092 -0.74891 0.453911
ar1     -1.011682  0.052106 -19.41568 0.000000
ar2     -0.188799  0.073950 -2.55307 0.010678
ar3     -0.080415  0.074949 -1.07294 0.283297
ar4     -0.049411  0.070068 -0.70519 0.480690
ar5     -0.021980  0.046900 -0.46865 0.639319
ma1      0.966466  0.006791 142.31782 0.000000
omega    0.170702  0.063023  2.70855 0.006758
alpha1    0.072330  0.019570  3.69599 0.000219
beta1    0.898139  0.020494 43.82409 0.000000

Robust Standard Errors:
            Estimate Std. Error   t value Pr(>|t|)
mu      -0.060731  0.091280 -0.66533 0.505841
ar1     -1.011682  0.064201 -15.75811 0.000000
ar2     -0.188799  0.076136 -2.47976 0.013147
ar3     -0.080415  0.095839 -0.83907 0.401431
ar4     -0.049411  0.084495 -0.58479 0.558691
ar5     -0.021980  0.064970 -0.33831 0.735132
ma1      0.966466  0.004038 239.31674 0.000000
omega    0.170702  0.085787  1.98984 0.046608
alpha1    0.072330  0.034213  2.11410 0.034507
beta1    0.898139  0.025868 34.71947 0.000000

LogLikelihood : -1128.912

Information Criteria
-----
Akaike        4.5739
Bayes         4.6585
Shibata       4.5732
Hannan-Quinn 4.6071

Weighted Ljung-Box Test on Standardized Residuals
-----
                           statistic p-value
Lag[1]                      0.09061  0.7634

```

Lag [2*(p+q) + (p+q)-1] [17] 6.90674 0.9999
 Lag [4*(p+q) + (p+q)-1] [29] 12.52177 0.8010
 d.o.f=6
 H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals

| | statistic | p-value |
|-----------------------------|-----------|---------|
| Lag [1] | 0.3167 | 0.5736 |
| Lag [2*(p+q) + (p+q)-1] [5] | 0.9209 | 0.8773 |
| Lag [4*(p+q) + (p+q)-1] [9] | 4.9988 | 0.4296 |

d.o.f=2

Weighted ARCH LM Tests

| | Statistic | Shape | Scale | P-Value |
|-------------|-----------|-------|-------|---------|
| ARCH Lag[3] | 0.1402 | 0.500 | 2.000 | 0.7081 |
| ARCH Lag[5] | 1.4545 | 1.440 | 1.667 | 0.6044 |
| ARCH Lag[7] | 5.5953 | 2.315 | 1.543 | 0.1709 |

Nyblom stability test

Joint Statistic: 3.255

Individual Statistics:

| | |
|--------|---------|
| mu | 0.53519 |
| ar1 | 0.20204 |
| ar2 | 0.15349 |
| ar3 | 0.11179 |
| ar4 | 0.07772 |
| ar5 | 0.12157 |
| ma1 | 0.05552 |
| omega | 0.08460 |
| alpha1 | 0.11291 |
| beta1 | 0.09846 |

Asymptotic Critical Values (10% 5% 1%)

Joint Statistic: 2.29 2.54 3.05

Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

| | t-value | prob | sig |
|--------------------|----------|--------|-----|
| Sign Bias | 0.369330 | 0.7120 | |
| Negative Sign Bias | 0.008559 | 0.9932 | |
| Positive Sign Bias | 1.020133 | 0.3082 | |
| Joint Effect | 2.650860 | 0.4486 | |

Adjusted Pearson Goodness-of-Fit Test:

| | group | statistic | p-value(g-1) |
|---|-------|-----------|--------------|
| 1 | 20 | 36.54 | 0.009059 |
| 2 | 30 | 49.59 | 0.009994 |
| 3 | 40 | 55.25 | 0.043974 |
| 4 | 50 | 78.10 | 0.005137 |

Elapsed time : 0.1988518

> show(garch_fit)

```
*-----*
*      GARCH Model Fit      *
*-----*
```

Conditional Variance Dynamics

GARCH Model : sGARCH(1,1)

Mean Model : ARFIMA(5,0,1)

Distribution : norm

Optimal Parameters

| | Estimate | Std. Error | t value | Pr(> t) |
|--------|-----------|------------|-----------|----------|
| mu | -0.060731 | 0.081092 | -0.74891 | 0.453911 |
| ar1 | -1.011682 | 0.052106 | -19.41568 | 0.000000 |
| ar2 | -0.188799 | 0.073950 | -2.55307 | 0.010678 |
| ar3 | -0.080415 | 0.074949 | -1.07294 | 0.283297 |
| ar4 | -0.049411 | 0.070068 | -0.70519 | 0.480690 |
| ar5 | -0.021980 | 0.046900 | -0.46865 | 0.639319 |
| ma1 | 0.966466 | 0.006791 | 142.31782 | 0.000000 |
| omega | 0.170702 | 0.063023 | 2.70855 | 0.006758 |
| alpha1 | 0.072330 | 0.019570 | 3.69599 | 0.000219 |
| beta1 | 0.898139 | 0.020494 | 43.82409 | 0.000000 |

Robust Standard Errors:

| | Estimate | Std. Error | t value | Pr(> t) |
|--------|-----------|------------|-----------|----------|
| mu | -0.060731 | 0.091280 | -0.66533 | 0.505841 |
| ar1 | -1.011682 | 0.064201 | -15.75811 | 0.000000 |
| ar2 | -0.188799 | 0.076136 | -2.47976 | 0.013147 |
| ar3 | -0.080415 | 0.095839 | -0.83907 | 0.401431 |
| ar4 | -0.049411 | 0.084495 | -0.58479 | 0.558691 |
| ar5 | -0.021980 | 0.064970 | -0.33831 | 0.735132 |
| ma1 | 0.966466 | 0.004038 | 239.31674 | 0.000000 |
| omega | 0.170702 | 0.085787 | 1.98984 | 0.046608 |
| alpha1 | 0.072330 | 0.034213 | 2.11410 | 0.034507 |
| beta1 | 0.898139 | 0.025868 | 34.71947 | 0.000000 |

LogLikelihood : -1128.912

Information Criteria

```
Akaike      4.5739
Bayes      4.6585
Shibata    4.5732
Hannan-Quinn 4.6071
```

Weighted Ljung-Box Test on Standardized Residuals

```
statistic p-value
Lag [1]          0.09061  0.7634
Lag [2 * (p+q) + (p+q)-1] [17] 6.90674  0.9999
Lag [4 * (p+q) + (p+q)-1] [29] 12.52177  0.8010
d.o.f=6
H0 : No serial correlation
```

Weighted Ljung-Box Test on Standardized Squared Residuals

```
statistic p-value
Lag [1]          0.3167   0.5736
Lag [2 * (p+q) + (p+q)-1] [5] 0.9209   0.8773
Lag [4 * (p+q) + (p+q)-1] [9] 4.9988   0.4296
d.o.f=2
```

Weighted ARCH LM Tests

| | Statistic | Shape | Scale | P-Value |
|-------------|-----------|-------|-------|---------|
| ARCH Lag[3] | 0.1402 | 0.500 | 2.000 | 0.7081 |
| ARCH Lag[5] | 1.4545 | 1.440 | 1.667 | 0.6044 |
| ARCH Lag[7] | 5.5953 | 2.315 | 1.543 | 0.1709 |

Nyblom stability test

Joint Statistic: 3.255

Individual Statistics:

```
mu      0.53519
ar1     0.20204
ar2     0.15349
ar3     0.11179
ar4     0.07772
ar5     0.12157
ma1     0.05552
omega   0.08460
alpha1  0.11291
beta1   0.09846
```

Asymptotic Critical Values (10% 5% 1%)

Joint Statistic: 2.29 2.54 3.05

Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

Adjusted Pearson Goodness-of-Fit Test:

| group | statistic | p-value(g-1) |
|-------|-----------|--------------|
| 1 | 20 | 36.54 |
| 2 | 30 | 49.59 |
| 3 | 40 | 55.25 |
| 4 | 50 | 78.10 |

Elapsed time : 0.1988518

Bài cá nhân - Nguyễn Thị Trà

Kết quả các mô hình cho chuỗi doanh thu PVS Mô hình lin-lin

```
Call:
lm(formula = train_ts ~ t_train)

Residuals:
    Min      1Q   Median      3Q     Max 
-3895.2 -852.7 -349.8  947.5 4126.8 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 6103.56    401.41 15.205 < 2e-16 ***
t_train     -37.10     12.25 -3.028 0.00377 **  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1482 on 54 degrees of freedom
Multiple R-squared:  0.1451, Adjusted R-squared:  0.1293 
F-statistic: 9.168 on 1 and 54 DF,  p-value: 0.003772
```

Mô hình lin-log

```
Call:
lm(formula = train_ts ~ log(t_train))

Residuals:
    Min      1Q   Median      3Q     Max 
-3945.6 -1001.9 -198.9  946.8 4432.8 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 6116.9     757.2  8.079 7.37e-11 ***
log(t_train) -347.8    236.4 -1.472    0.147  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1572 on 54 degrees of freedom
Multiple R-squared:  0.03856, Adjusted R-squared:  0.02075 
F-statistic: 2.166 on 1 and 54 DF,  p-value: 0.1469
```

Mô hình log-lin

```

Call:
lm(formula = log(train_ts) ~ t_train)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.98057 -0.13226 -0.01931  0.21281  0.61940 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 8.670377   0.080588 107.59 < 2e-16 ***
t_train     -0.006739   0.002460   -2.74  0.00831 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2975 on 54 degrees of freedom
Multiple R-squared:  0.122,    Adjusted R-squared:  0.1058 
F-statistic: 7.507 on 1 and 54 DF,  p-value: 0.008313

```

Mô hình log-log

```

Call:
lm(formula = log(train_ts) ~ log(t_train))

Residuals:
    Min      1Q  Median      3Q     Max 
-0.95963 -0.20312 -0.00563  0.22203  0.67630 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 8.64269   0.15120  57.162 <2e-16 ***
log(t_train) -0.05341   0.04720   -1.132   0.263  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3138 on 54 degrees of freedom
Multiple R-squared:  0.02316,  Adjusted R-squared:  0.005074 
F-statistic: 1.28 on 1 and 54 DF,  p-value: 0.2628

```

Mô hình bậc 2 của trend

```

Call:
lm(formula = train_ts ~ t_train + I(t_train^2))

Residuals:
    Min      1Q  Median      3Q     Max 
-3622.2 -846.5 -328.0  852.7  4032.5 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 5799.6468   619.2852   9.365 7.92e-13 ***
t_train     -5.6566    50.1277  -0.113   0.911  
I(t_train^2) -0.5516    0.8525  -0.647   0.520  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1490 on 53 degrees of freedom
Multiple R-squared:  0.1518,  Adjusted R-squared:  0.1198 
F-statistic: 4.744 on 2 and 53 DF,  p-value: 0.01273

```

Mô hình biến giả mùa vụ

```

Call:
lm(formula = train_ts ~ s2 + s3 + s4)

Residuals:
    Min      1Q  Median      3Q     Max 
-2413.7 -1134.5 -258.7 1010.1 4147.5 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 3914.9     397.0   9.861 1.68e-13 ***
s2          1562.2     561.5   2.782   0.0075 **  
s3          1463.2     561.5   2.606   0.0119 *   
s4          1500.5     561.5   2.672   0.0100 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1485 on 52 degrees of freedom
Multiple R-squared:  0.1728, Adjusted R-squared:  0.1251 
F-statistic:  3.62 on 3 and 52 DF,  p-value: 0.01894

```

Mô hình tuyến tính + mùa vụ, dạng cộng

```

Call:
lm(formula = train_ts ~ t_train + s2 + s3 + s4)

Residuals:
    Min      1Q  Median      3Q     Max 
-2767.7 -865.8   17.1   726.1 3753.6 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 4978.40     468.55 10.625 1.57e-14 ***
t_train     -39.39     11.14 -3.536 0.000875 *** 
s2          1601.64     508.20  3.152 0.002718 **  
s3          1541.97     508.56  3.032 0.003813 **  
s4          1618.62     509.17  3.179 0.002513 **  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1344 on 51 degrees of freedom
Multiple R-squared:  0.3356, Adjusted R-squared:  0.2835 
F-statistic: 6.442 on 4 and 51 DF,  p-value: 0.0002829

```

Mô hình tuyến tính + mùa vụ, dạng nhân

```

Call:
lm(formula = train_ts ~ t_train + I(t_train * s2) + I(t_train *
s3) + I(t_train * s4))

Residuals:
    Min      1Q  Median      3Q     Max 
-3899.2 -674.9 -104.1  775.2 4026.9 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 6134.50     389.70 15.742 < 2e-16 ***
t_train     -64.04      16.20 -3.953 0.000239 ***  
I(t_train * s2) 30.81      17.05  1.807 0.076669 .  
I(t_train * s3) 30.57      16.84  1.816 0.075319 :  
I(t_train * s4) 39.94      16.64  2.400 0.020069 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1438 on 51 degrees of freedom
Multiple R-squared:  0.24,   Adjusted R-squared:  0.1804 
F-statistic: 4.027 on 4 and 51 DF,  p-value: 0.006501

```

Mô hình không tuyến tính + mùa vụ, dạng cộng

```

Call:
lm(formula = log(train_ts) ~ t_train + s2 + s3 + s4)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.74374 -0.16620  0.03214  0.15849  0.55082 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 8.434027  0.092694 90.988 < 2e-16 ***
t_train     -0.007223  0.002204 -3.278 0.00189 **  
s2          0.340363  0.100536  3.385 0.00137 **  
s3          0.314129  0.100609  3.122 0.00295 **  
s4          0.346116  0.100729  3.436 0.00118 **  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2659 on 51 degrees of freedom
Multiple R-squared:  0.3375,   Adjusted R-squared:  0.2855 
F-statistic: 6.495 on 4 and 51 DF,  p-value: 0.0002649

```

Mô hình không tuyến tính + mùa vụ, dạng nhân

```

Call:
lm(formula = log(train_ts) ~ t_train + I(t_train * s2) +
I(t_train *
s3) + I(t_train * s4))

Residuals:
    Min      1Q  Median      3Q     Max 
-0.98135 -0.14007  0.02108  0.18074  0.60106 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 8.676903   0.077576 111.850 < 2e-16 ***
t_train     -0.012493   0.003225 -3.874 0.000307 ***  
I(t_train * s2) 0.006749   0.003394  1.988 0.052171 .  
I(t_train * s3) 0.006376   0.003352  1.902 0.062802 .  
I(t_train * s4) 0.008534   0.003312  2.576 0.012927 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2862 on 51 degrees of freedom
Multiple R-squared:  0.2326,    Adjusted R-squared:  0.1724 
F-statistic: 3.865 on 4 and 51 DF,  p-value: 0.008102

```

Mô hình Holt-Winter có mùa vụ, dạng cộng

Holt-Winters exponential smoothing with trend and additive seasonal component.

```

Call:
HoltWinters(x = train_ts, seasonal = "a")

Smoothing parameters:
alpha: 0.3888657
beta : 0.1558883
gamma: 0.4155563

Coefficients:
[,1]
a 5547.0896
b 129.8870
s1 -1292.5699
s2 -430.3554
s3 -438.6385
s4 575.6250

```

Mô hình Holt-Winter có mùa vụ, dạng nhân

Holt-Winters exponential smoothing with trend and multiplicative seasonal component.

Call:
HoltWinters(x = train_ts, seasonal = "m")

Smoothing parameters:

alpha: 0.4654209
beta : 0.1446132
gamma: 0.5517191

Coefficients:

| | [,1] |
|----|--------------|
| a | 5908.7863119 |
| b | 134.0998790 |
| s1 | 0.6991583 |
| s2 | 0.8480018 |
| s3 | 0.8457521 |
| s4 | 1.0799589 |

Kết quả các mô hình ARIMA cho chuỗi giá đóng cửa PVS ARIMA(6,1,1)

Series: price
ARIMA(6,1,1)

Coefficients:

| | ar1 | ar2 | ar3 | ar4 | ar5 | ar6 | ma1 |
|------|---------|---------|---------|---------|--------|---------|--------|
| - | -0.2423 | -0.0658 | -0.0017 | -0.0486 | 0.1075 | -0.0722 | 0.1482 |
| s.e. | 0.7293 | 0.0896 | 0.0566 | 0.0458 | 0.0556 | 0.1082 | 0.7331 |

$\sigma^2 = 482158$: log likelihood = -3961.61
AIC=7939.21 AICc=7939.51 BIC=7972.9

Training set error measures:

| | ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
|--------------|---------|----------|----------|------------|----------|-----------|--------------|
| Training set | 28.6064 | 688.7875 | 482.6052 | 0.08229468 | 1.411958 | 0.9773109 | -0.001889557 |

ARIMA(6,1,0)

Series: price
ARIMA(6,1,0)

Coefficients:

| | ar1 | ar2 | ar3 | ar4 | ar5 |
|------|---------|---------|--------|---------|--------|
| - | -0.0953 | -0.0501 | 0.0051 | -0.0494 | 0.1142 |
| s.e. | 0.0446 | 0.0445 | 0.0444 | 0.0444 | 0.0443 |
| ar6 | | | | | |
| - | -0.0907 | | | | |
| s.e. | 0.0444 | | | | |

$\sigma^2 = 481241$: log likelihood = -3961.64
AIC=7937.28 AICc=7937.5 BIC=7966.75

Training set error measures:

| | ME | RMSE | MAE | MPE |
|--------------|----------|-----------|--------------|------------|
| Training set | 28.95141 | 688.8327 | 482.3437 | 0.08334391 |
| | MAPE | MASE | ACF1 | |
| Training set | 1.411072 | 0.9767815 | -0.001151392 | |

ARIMA(5,1,5)

```

Series: price
ARIMA(5,1,5)

Coefficients:
ar1      ar2      ar3      ar4      ar5      ma1      ma2      ma3
-0.5439 -0.3199 -0.0221  0.5186  0.5703  0.4528  0.2348 -0.0163
s.e.    0.2330  0.1853  0.2141  0.1718  0.1469  0.2460  0.1637  0.1842
ma4      ma5
-0.6150 -0.4632
s.e.    0.1462  0.1741

sigma^2 = 481123: log likelihood = -3959.64
AIC=7941.27   AICc=7941.81   BIC=7987.59

Training set error measures:
          ME      RMSE      MAE      MPE      MAPE      MASE
Training set 33.2932 685.9422 480.8599 0.0970128 1.409417 0.9737766
          ACF1
Training set -0.006428796

```

ARIMA(5,1,1)

```

Series: price
ARIMA(5,1,1)

Coefficients:
ar1      ar2      ar3      ar4      ar5      ma1
-0.6131 -0.1038 -0.0184 -0.0451  0.0997  0.5188
s.e.    0.1980  0.0567  0.0533  0.0522  0.0531  0.1964

sigma^2 = 481553: log likelihood = -3961.8
AIC=7937.6   AICc=7937.82   BIC=7967.07

Training set error measures:
          ME      RMSE      MAE      MPE      MAPE      MASE
Training set 27.48705 689.0554 483.7859 0.07891262 1.4155 0.979702
          ACF1
Training set -0.0004477986

```

ARIMA(1,1,6)

```

Series: price
ARIMA(1,1,6)

Coefficients:
ar1      ma1      ma2      ma3      ma4      ma5      ma6
-0.1291  0.0336 -0.0567  0.0081 -0.0447  0.1040 -0.0965
s.e.    0.3849  0.3837  0.0525  0.0494  0.0459  0.0459  0.0701

sigma^2 = 482766: log likelihood = -3961.91
AIC=7939.83   AICc=7940.12   BIC=7973.51

Training set error measures:
          ME      RMSE      MAE      MPE      MAPE      MASE
Training set 29.5606 689.2215 480.9918 0.08515805 1.406591 0.9740437
          ACF1
Training set -0.001921513

```

ARIMA(1,1,5)

```

Series: price
ARIMA(1,1,5)

Coefficients:
      ar1      ma1      ma2      ma3      ma4      ma5
     -0.4768   0.3848  -0.0855  -0.0067  -0.0412   0.1040
s.e.   0.1858   0.1853   0.0494   0.0508   0.0482   0.0491

sigma^2 = 483235: log likelihood = -3962.65
AIC=7939.3   AICc=7939.53   BIC=7968.78

```

```

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE
Training set 27.03712 690.2583 484.3002 0.07747535 1.416135 0.9807436
          ACF1
Training set -0.004464064

```

ARIMA(1,1,1)

```

Series: price
ARIMA(1,1,1)

Coefficients:
      ar1      ma1
     0.2174  -0.3332
s.e.  0.3259   0.3147

sigma^2 = 490882: log likelihood = -3968.52
AIC=7943.04   AICc=7943.09   BIC=7955.68

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE
Training set 29.06046 698.5202 486.2374 0.08310895 1.418569 0.9846665
          ACF1
Training set -0.0009654295

```

Kết quả các mô hình ARIMA cho chuỗi log-return giá PVS ARIMA(5,0,6)

```

Series: logreturn
ARIMA(5,0,6) with zero mean

Coefficients:
Warning: NaNs produced      ar1      ar2      ar3      ar4      ar5      ma1      ma2      ma3      ma4
      -0.2196   0.2934  -0.3936   0.2234   0.9155   0.1197  -0.3632   0.4623  -0.2846
s.e.       NaN       NaN    0.0061       NaN       NaN    0.0328       NaN    0.0234   0.0130
          ma5      ma6
         -0.8901   0.1242
s.e.       NaN    0.0406

sigma^2 = 4.013: log likelihood = -1048.12
AIC=2120.24   AICc=2120.88   BIC=2170.76

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.09766518 1.980971 1.398331  NaN    Inf  0.6256769 -0.01076308

```

ARIMA(5,0,5)

```

Series: logreturn
ARIMA(5,0,5) with zero mean

Coefficients:
      ar1     ar2     ar3     ar4     ar5     ma1     ma2     ma3
    -0.0916  -0.0124  -0.3872  0.1564  0.6544  -0.0204  0.0055  0.3951
  s.e.   0.2324   0.1618   0.1447  0.2010  0.1568   0.2482  0.1639  0.1341
      ma4     ma5
    -0.2765  -0.5678
  s.e.   0.2029   0.1700

sigma^2 = 4.021: log likelihood = -1048.91
AIC=2119.82  AICC=2120.36  BIC=2166.14

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.1158933 1.984966 1.401614 NaN Inf 0.6271458 -0.005987866

```

ARIMA(5,0,1)

```

Series: logreturn
ARIMA(5,0,1) with zero mean

Coefficients:
      ar1     ar2     ar3     ar4     ar5     ma1
    -0.5882  -0.0779  -0.0152  -0.0371  0.1012  0.4781
  s.e.   0.2277   0.0589   0.0520   0.0518  0.0525  0.2270

sigma^2 = 4.046: log likelihood = -1051.71
AIC=2117.41  AICC=2117.64  BIC=2146.89

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.09953793 1.999281 1.414918 NaN Inf 0.6330988 -0.001559333

```

ARIMA(1,0,6)

```

Series: logreturn
ARIMA(1,0,6) with zero mean

Coefficients:
      ar1     ma1     ma2     ma3     ma4     ma5     ma6
    -0.1198  0.0081  -0.0268  -0.0009  -0.0327  0.1049  -0.0957
  s.e.   0.4801  0.4795   0.0627   0.0460   0.0455  0.0446   0.0753

sigma^2 = 4.054: log likelihood = -1051.67
AIC=2119.34  AICC=2119.63  BIC=2153.02

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.1063846 1.999133 1.408111 NaN Inf 0.6300531 -0.002520638

```

ARIMA(1,0,5)

```

Series: logreturn
ARIMA(1,0,5) with zero mean

Coefficients:
ar1      ma1      ma2      ma3      ma4      ma5
-0.5186  0.4110  -0.0642  -0.0058  -0.0343  0.1051
s.e.    0.1967  0.1963   0.0516   0.0503   0.0494  0.0488

sigma^2 = 4.056: log likelihood = -1052.29
AIC=2118.58  AICc=2118.81  BIC=2148.05

Training set error measures:
          ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.09777553 2.001661 1.416343 NaN     Inf 0.6337363 -0.004958689

```

ARIMA(1,0,1)

```

Series: logreturn
ARIMA(1,0,1) with zero mean

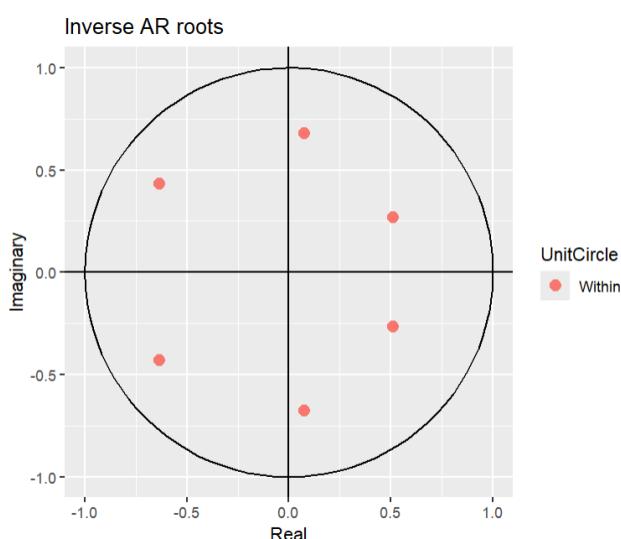
Coefficients:
ar1      ma1
0.0500 -0.1754
s.e.    0.3722 0.3668

sigma^2 = 4.106: log likelihood = -1057.31
AIC=2120.62  AICc=2120.67  BIC=2133.25

Training set error measures:
          ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.1045716 2.022148 1.421129 NaN     Inf 0.6358777 -0.003218987

```

**Kiểm định về mô hình
ARIMA(6,1,0)**

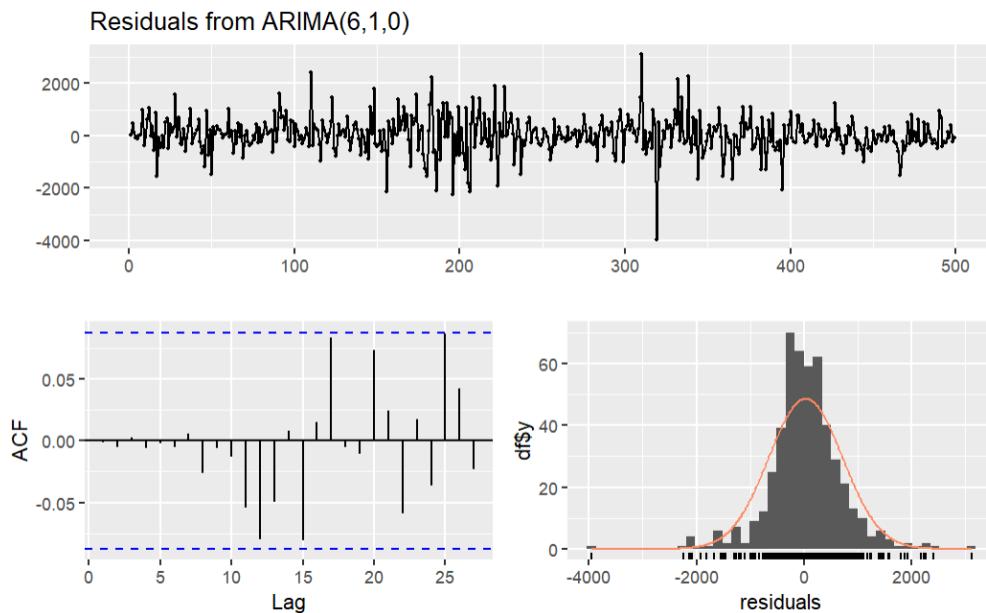


Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị.

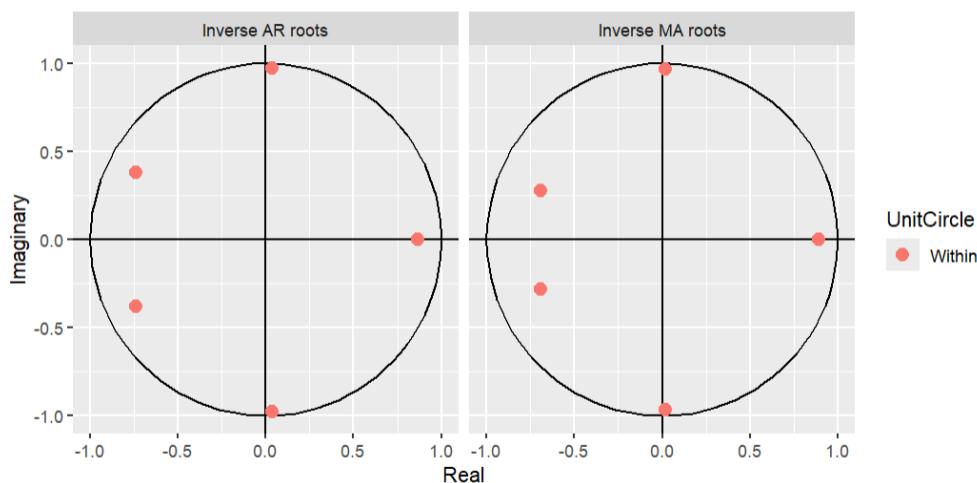
Ljung-Box test

```
data: Residuals from ARIMA(6,1,0)
Q* = 0.51779, df = 4, p-value = 0.9717
```

Model df: 6. Total lags used: 10



$p-value = 0.9717 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

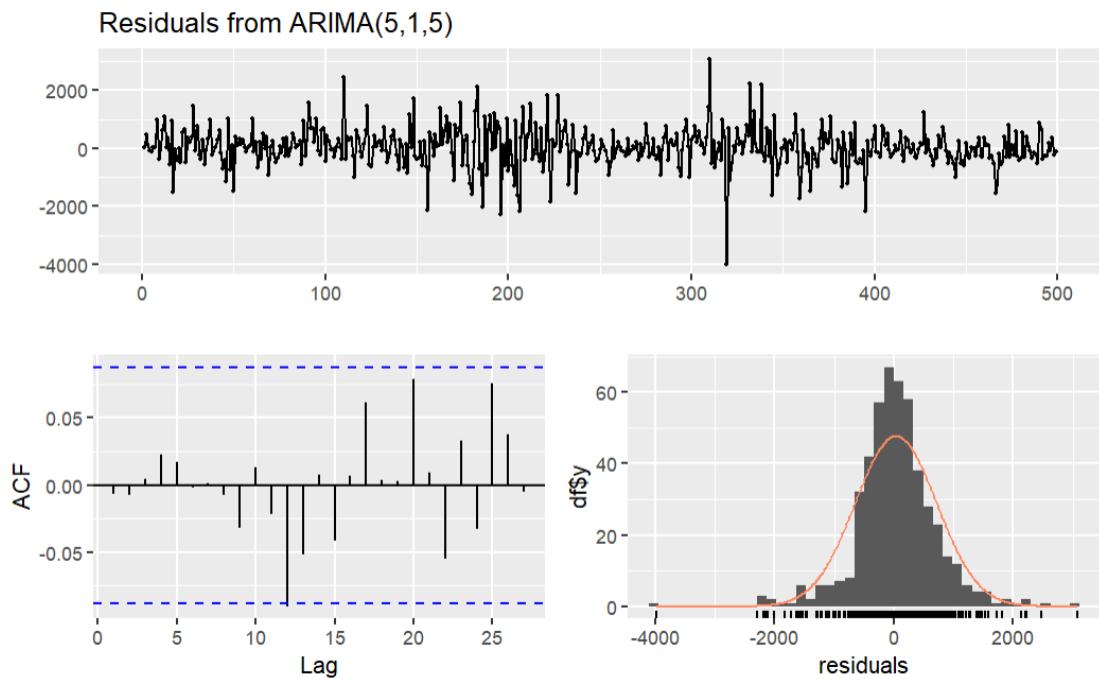
ARIMA(5,1,5)

Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị.

Ljung-Box test

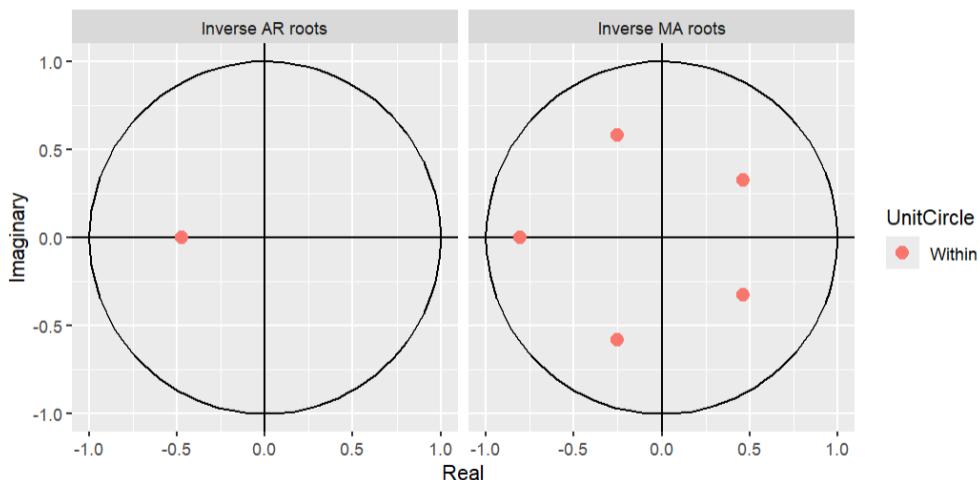
```
data: Residuals from ARIMA(5,1,5)
Q* = 6.9058, df = 3, p-value = 0.07496
```

Model df: 10. Total lags used: 13



$p - value = 0.07496 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

ARIMA(1,1,5)

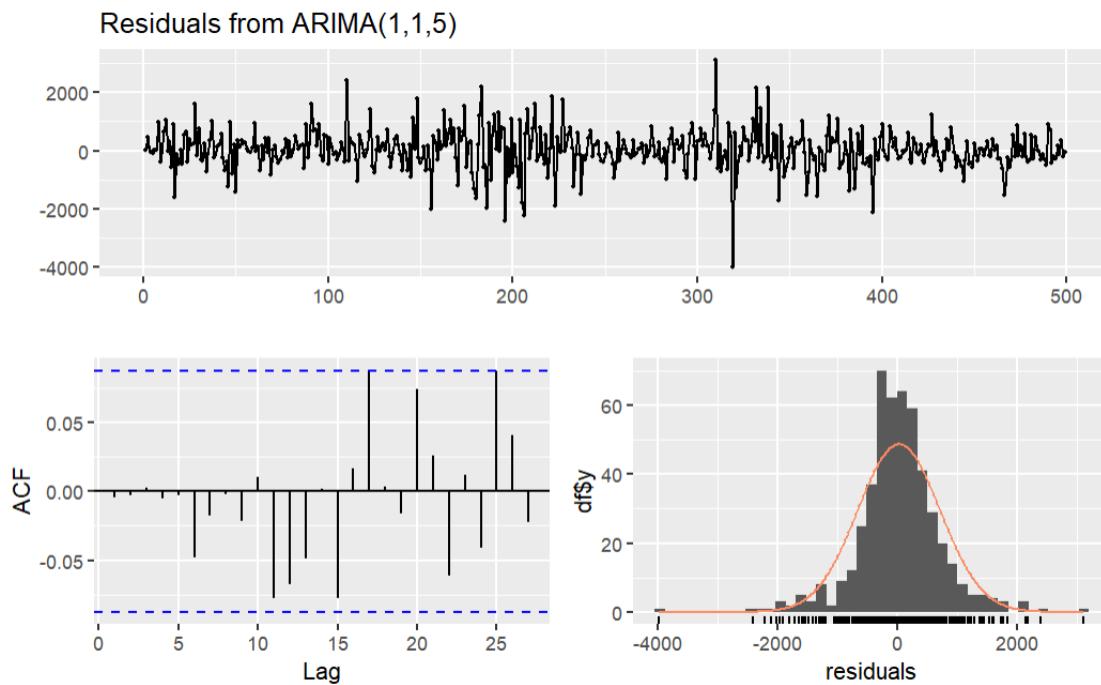


Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị.

Ljung-Box test

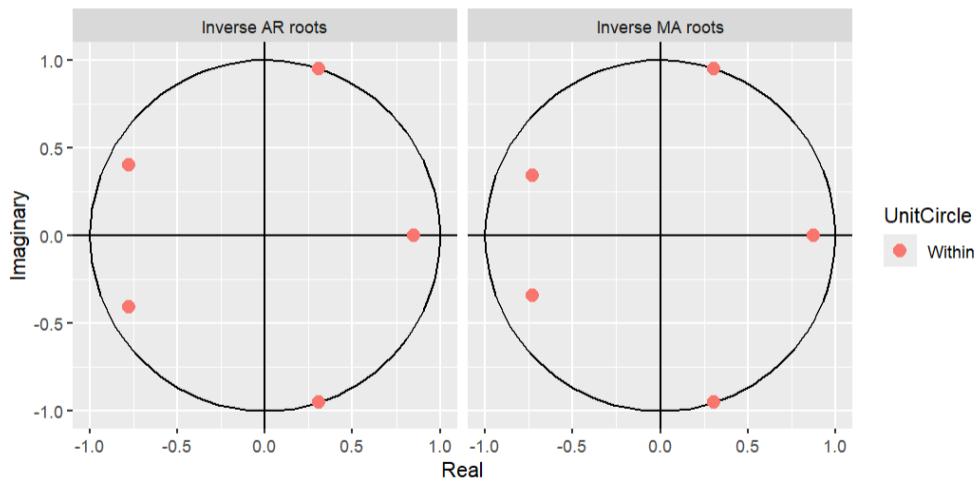
```
data: Residuals from ARIMA(1,1,5)
Q* = 1.6865, df = 4, p-value = 0.7932

Model df: 6. Total lags used: 10
```



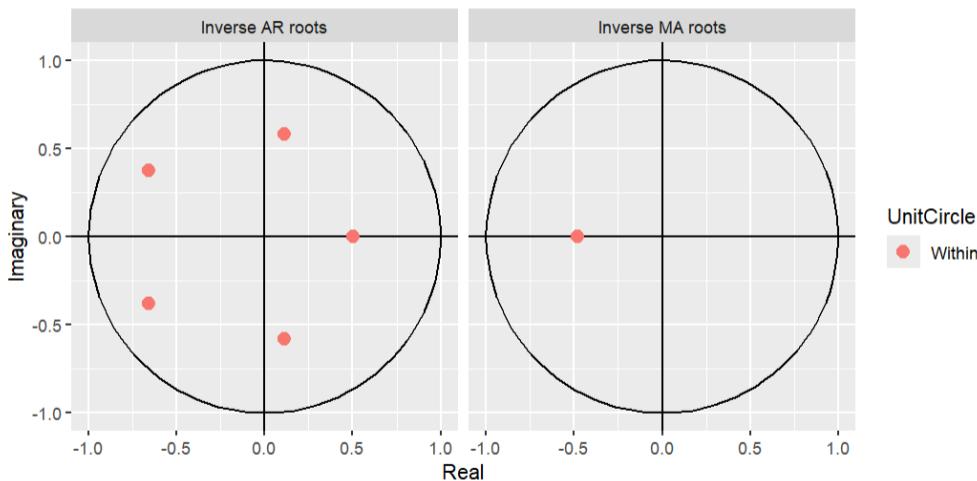
$p - value = 0.7932 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

ARIMA(5,0,5)



Nghiệm nghịch đảo đều nằm trên đường tròn đơn vị.

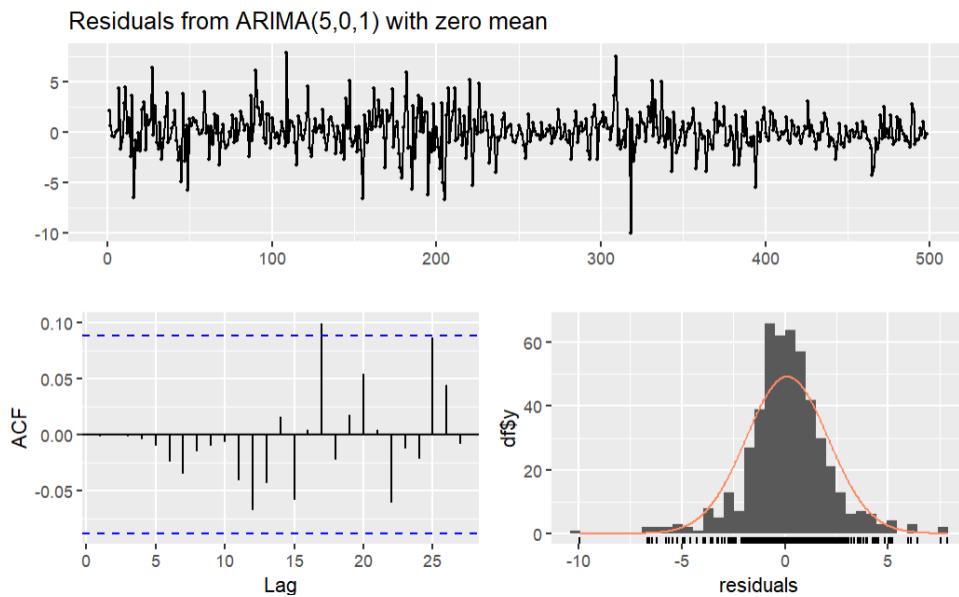
ARIMA(5,0,1)



Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị.

Box-Ljung test

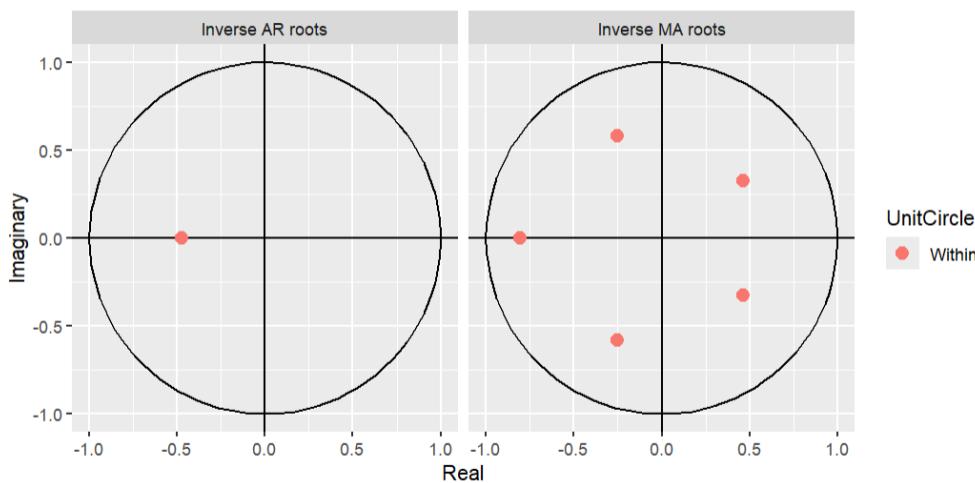
```
data: residuals(log.arima501)
X-squared = 7.123, df = 16, p-value = 0.9708
```



Đồ thị ACF có bậc 16 vượt ngoài ngưỡng tin cậy.

Tuy nhiên $p - value = 0.9708 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

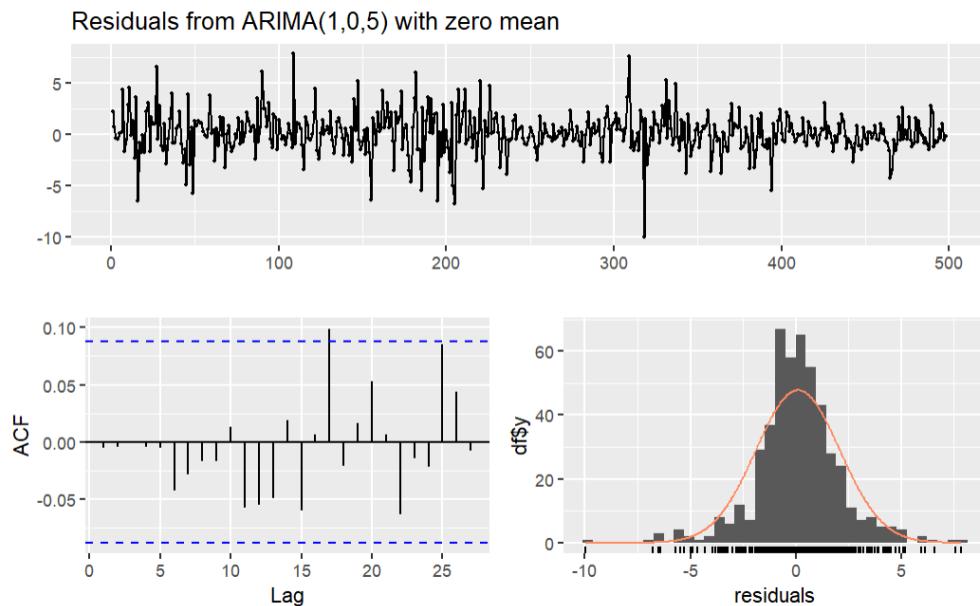
ARIMA(1,0,5)



Nghiệm nghịch đảo đều thuộc vòng tròn đơn vị.

Box-Ljung test

```
data: residuals(log.arima105)
X-squared = 8.2548, df = 16, p-value = 0.9409
```



Đồ thị ACF có bậc 16 vượt ngoài ngưỡng tin cậy.

Tuy nhiên, $p - value = 0.9409 > 0.05 \rightarrow$ Chưa bác bỏ $H_0 \rightarrow$ Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng.

Kết quả chuỗi log-return và giá chuyển đổi tương ứng ARIMA(1,0,5)

Chuỗi log-return dự báo

```
Time Series:
Start = 499
End = 508
Frequency = 1
[1] 0.149243444 -0.131101752 0.193054318 -0.147287127 0.052619281 -0.027289798 0.014153236
[8] -0.007340255 0.003806857 -0.001974340
```

RMSE = 1.649092

Chuỗi giá chuyển đổi từ log-return

```
[1] 33950.63 33906.15 33971.67 33921.67 33939.53 33930.27 33935.07 33932.58 33933.87 33933.20
```

ARIMA(5,0,1)

Chuỗi log-return dự báo

```
Time Series:
Start = 499
End = 508
Frequency = 1
[1] 0.21284127 -0.15579514 0.19044433 -0.15154493 0.03898291 0.01328820 -0.03137156 0.04171381
[9] -0.03907413 0.02366375
```

RMSE = 1.64071

Chuỗi giá chuyển đổi từ log-return

```
[1] 33972.23 33919.34 33984.00 33932.54 33945.77 33950.28 33939.63 33953.79 33940.53 33948.56
```

Kết quả mô hình GARCH cho chuỗi log-return giá PVS

```
Title:
GARCH Modelling

Call:
garchFit(formula = ~arma(1, 5) + garch(1, 1), data = logreturn,
trace = FALSE)

Mean and Variance Equation:
data ~ arma(1, 5) + garch(1, 1)
<environment: 0x000001fd82dad240>
[data = logreturn]

Conditional Distribution:
norm

Coefficient(s):
      mu        ar1        ma1        ma2        ma3
 0.0123888 -0.5523797  0.4988676 -0.0600985  0.0071682
      ma4        ma5       omega     alpha1      beta1
-0.0010343  0.0939124  0.0994398  0.0684009  0.9083132

Std. Errors:
based on Hessian

Error Analysis:
    Estimate Std. Error t value Pr(>|t|)
mu      0.012389   0.124891   0.099 0.920981
ar1     -0.552380   0.253947  -2.175 0.029617 *
ma1      0.498868   0.253446   1.968 0.049029 *
ma2     -0.060099   0.057185  -1.051 0.293283
ma3      0.007168   0.059389   0.121 0.903928
ma4     -0.001034   0.059041  -0.018 0.986023
ma5      0.093912   0.055589   1.689 0.091139 .
omega    0.099440   0.045497   2.186 0.028844 *
alpha1   0.068401   0.019176   3.567 0.000361 ***
beta1    0.908313   0.021239  42.766 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
-1033.251   normalized: -2.074802

Description:
Mon Apr 14 23:17:47 2025 by user: caboiifvd
```

Standardised Residuals Tests:

| | | | Statistic | p-value |
|-------------------|-----|-------|-------------|--------------|
| Jarque-Bera Test | R | Chi^2 | 192.6872501 | 0.000000e+00 |
| Shapiro-Wilk Test | R | W | 0.9603742 | 2.547445e-10 |
| Ljung-Box Test | R | Q(10) | 1.6732583 | 9.982866e-01 |
| Ljung-Box Test | R | Q(15) | 3.6993794 | 9.985673e-01 |
| Ljung-Box Test | R | Q(20) | 8.4481346 | 9.884630e-01 |
| Ljung-Box Test | R^2 | Q(10) | 13.6445079 | 1.898328e-01 |
| Ljung-Box Test | R^2 | Q(15) | 18.5392465 | 2.353770e-01 |
| Ljung-Box Test | R^2 | Q(20) | 23.3836044 | 2.703536e-01 |
| LM Arch Test | R | TR^2 | 13.5100572 | 3.330819e-01 |

Information Criterion Statistics:

| AIC | BIC | SIC | HQIC |
|----------|----------|----------|----------|
| 4.189764 | 4.274314 | 4.188978 | 4.222947 |

Bài nhóm

Kiểm định ADF cho chuỗi lợi suất danh mục, có xu thế:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression trend
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -4.6718 -0.5811 -0.0179  0.6224  4.6430 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 0.1557984  0.0976620   1.595   0.111    
## z.lag.1     -1.0975592  0.0666685  -16.463  <2e-16 ***  
## tt          -0.0003558  0.0003385   -1.051   0.294    
## z.diff.lag   0.0044543  0.0451145    0.099   0.921    
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 1.077 on 492 degrees of freedom
## Multiple R-squared:  0.5458, Adjusted R-squared:  0.543  
## F-statistic: 197.1 on 3 and 492 DF,  p-value: < 2.2e-16 
## 
## 
## Value of test-statistic is: -16.4629 90.3475 135.514
## 
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.98 -3.42 -3.13
## phi2  6.15  4.71  4.05
## phi3  8.34  6.30  5.36

```

Bảng 5.39: Kiểm định ADF cho chuỗi lợi suất suất danh mục, có xu thế

Kiểm định ADF cho chuỗi lợi suất danh mục, có hệ số chặn:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression drift
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
## 
## Residuals:
##      Min        1Q    Median        3Q       Max
## -4.5944 -0.5729 -0.0245  0.6190  4.6259
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 0.066738  0.048574   1.374   0.170    
## z.lag.1     -1.093142  0.066543 -16.428 <2e-16 *** 
## z.diff.lag   0.002195  0.045068   0.049   0.961    
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 1.077 on 493 degrees of freedom
## Multiple R-squared:  0.5448, Adjusted R-squared:  0.5429 
## F-statistic:  295 on 2 and 493 DF,  p-value: < 2.2e-16  
## 
## 
## Value of test-statistic is: -16.4276 134.9401
## 
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phil  6.47  4.61  3.79

```

Bảng 5.40: Kiểm định ADF cho chuỗi lợi suất danh mục, có hệ số chặn

Kiểm định ADF cho chuỗi lợi suất danh mục, không có hệ số chặn:

```

## 
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
## 
## Test regression none
## 
## 
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -4.5175 -0.5055  0.0429  0.6817  4.6828 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## z.lag.1    -1.084831   0.066327 -16.356   <2e-16 ***
## z.diff.lag -0.001998   0.045005  -0.044    0.965  
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 1.078 on 494 degrees of freedom
## Multiple R-squared:  0.543, Adjusted R-squared:  0.5412 
## F-statistic: 293.5 on 2 and 494 DF,  p-value: < 2.2e-16 
## 
## 
## Value of test-statistic is: -16.3558
## 
## Critical values for test statistics:
##      1pct  5pct 10pct 
## tau1 -2.58 -1.95 -1.62

```

Bảng 5.41: Kiểm định ADF cho chuỗi lợi suất danh mục, không có hệ số chặn

Mô hình ARIMA(1,0,1) cho chuỗi lợi suất danh mục:

```

## Series: newport
## ARIMA(1,0,1) with non-zero mean
## 
## Coefficients:
##             ar1      ma1      mean
##           0.4654  -0.5503  0.0637
## s.e.  0.4110   0.3887  0.0405
## 
## sigma^2 = 1.158: log likelihood = -741.62
## AIC=1491.23  AICc=1491.32  BIC=1508.08
## 
## Training set error measures:
##                   ME       RMSE       MAE       MPE       MAPE
## MASE
## Training set 0.0008555316 1.072765 0.8006279 53.89191 160.7015
## 0.6654825
##                   ACF1
## Training set -0.00804703

```

Bảng 5.42: Mô hình ARIMA(1,0,1) cho chuỗi lợi suất danh mục

Mô hình ARIMA(1,0,0) cho chuỗi lợi suất danh mục:

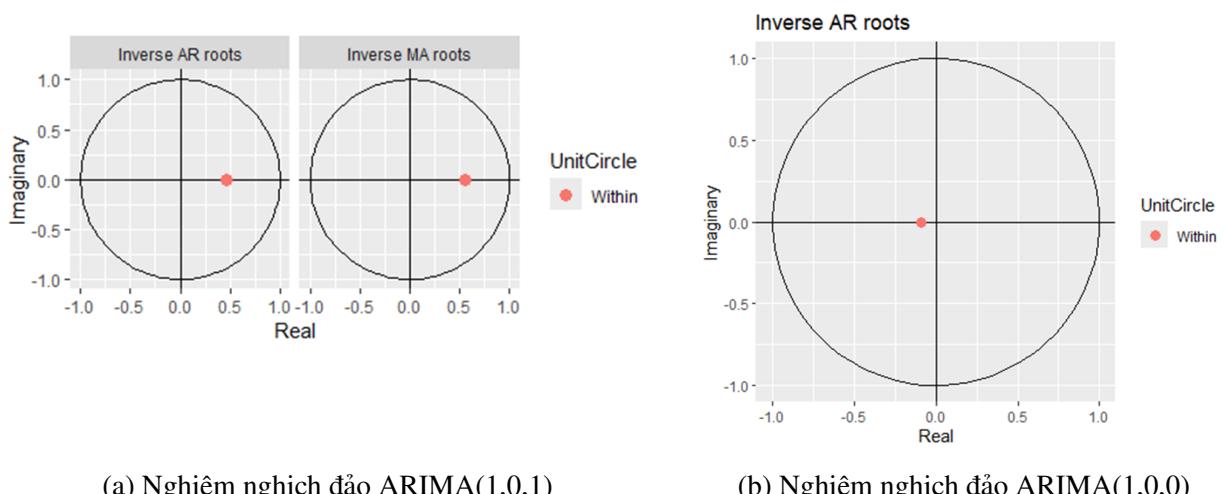
```

## Series: newport
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##             ar1      mean
##            -0.0894  0.0641
## s.e.      0.0447  0.0442
##
## sigma^2 = 1.157: log likelihood = -741.89
## AIC=1489.78   AICc=1489.83   BIC=1502.41
##
## Training set error measures:
##                         ME      RMSE      MAE      MPE      MAPE
MASE
## Training set 0.0001265784 1.073356 0.8018643 67.63379 145.5155
0.6665101
##                      ACF1
## Training set -0.0002868293

```

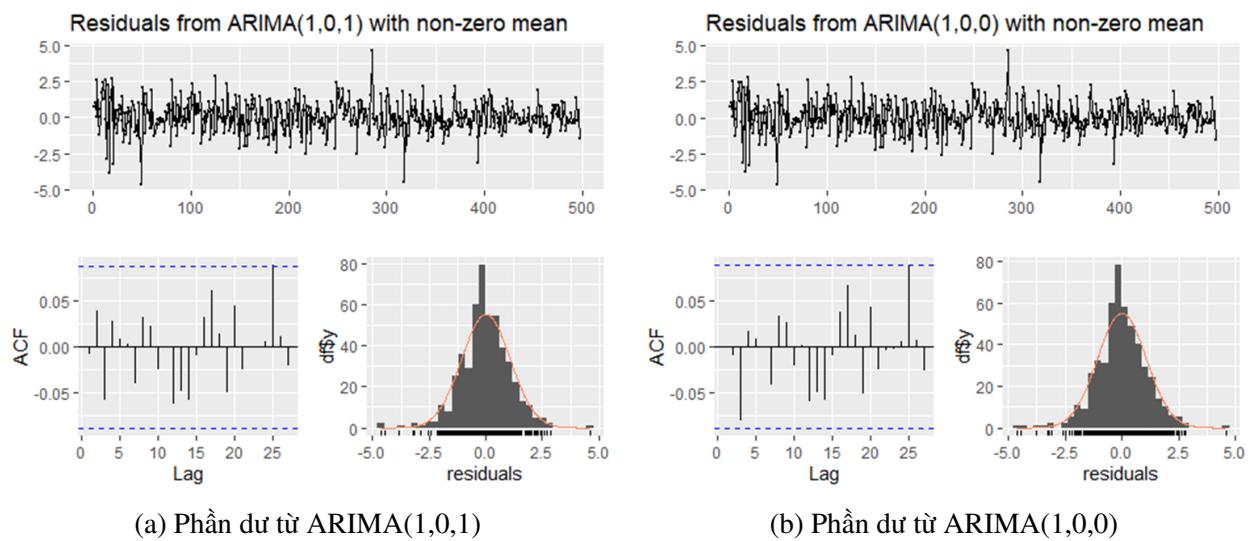
Bảng 5.43: Mô hình ARIMA(1,0,0) cho chuỗi lợi suất danh mục

Nghiệm nghịch đảo đều nằm trong vòng tròn đơn vị:



Hình 5.6: Xem tính dừng qua nghiệm nghịch đảo

Không tồn tại hiện tượng tự tương quan, phần dư nhiễu trắng:



Hình 5.7: Kiểm định tính nhiễu trắng của phân dư

Kiểm định đồng tích hợp:

```

## #####
## # Johansen-Procedure #
## #####
##
## Test type: trace statistic , with linear trend
##
## Eigenvalues (lambda):
## [1] 3.282506e-02 1.896249e-02 1.380096e-02 2.229126e-07
##
## Values of teststatistic and critical values of test:
##
##          test 10pct   5pct   1pct
## r <= 3 | 0.00  6.50  8.18 11.65
## r <= 2 | 6.91 15.66 17.95 23.52
## r <= 1 | 16.42 28.71 31.52 37.22
## r = 0 | 33.01 45.23 48.28 55.43
##
## Eigenvectors, normalised to first column:
## (These are the cointegration relations)
##
##           VCB.12      CTG.12      PVC.12      PVS.12
## VCB.12  1.00000000  1.000000  1.000000  1.000000
## CTG.12 -0.73285514 -12.140287  1.667047 18.439072
## PVC.12 -0.09730136 -34.880706  8.546040 -29.360311
## PVS.12 -0.37877650  9.957891 -9.563819 -3.707585
##
## Weights W:
## (This is the loading matrix)
##
##           VCB.12      CTG.12      PVC.12      PVS.12
## VCB.d -0.044100253  0.0004456147 0.0017126057 -2.106698e-06
## CTG.d  0.005338205  0.0009327450 0.0002198756 -2.359742e-06
## PVC.d  0.003866150  0.0010378336 0.0005624870  2.567222e-07
## PVS.d  0.016778709  0.0008944560 0.0017992771 -2.727989e-07

```

Bảng 5.44: Kết quả kiểm định đồng tích hợp

```

VAR Estimation Results:
=====
Endogenous variables: rVCB, rCTG, rPVC, rPVS
Deterministic variables: none
Sample size: 497
Log Likelihood: -3601.626
Roots of the characteristic polynomial:
0.1446 0.1446 0.06997 0.06997
Call:
VAR(y = data1, p = 1, type = "none")

Estimation results for equation rVCB:
=====
rVCB = rVCB.l1 + rCTG.l1 + rPVC.l1 + rPVS.l1

      Estimate Std. Error t value Pr(>|t|) 
rVCB.l1 -0.05948   0.04768 -1.247   0.213 
rCTG.l1  0.03083   0.03835  0.804   0.422 
rPVC.l1  0.03812   0.03593  1.061   0.289 
rPVS.l1 -0.05630   0.04587 -1.228   0.220 

Residual standard error: 1.204 on 493 degrees of freedom
Multiple R-Squared: 0.006762, Adjusted R-squared: -0.001297
F-statistic: 0.8391 on 4 and 493 DF, p-value: 0.5008

Estimation results for equation rCTG:
=====
rCTG = rVCB.l1 + rCTG.l1 + rPVC.l1 + rPVS.l1

      Estimate Std. Error t value Pr(>|t|) 
rVCB.l1 -0.04300   0.06383 -0.674   0.5009 
rCTG.l1 -0.11531   0.05135 -2.246   0.0252 *  
rPVC.l1  0.01476   0.04810  0.307   0.7590 
rPVS.l1 -0.05029   0.06140 -0.819   0.4132 

---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.612 on 493 degrees of freedom
Multiple R-Squared: 0.02373, Adjusted R-squared: 0.01581
F-statistic: 2.996 on 4 and 493 DF, p-value: 0.01839

```

Kết quả dự báo từ mô hình VAR(1)

```
Estimation results for equation rPVC:
=====
rPVC = rVCB.l1 + rCTG.l1 + rPVC.l1 + rPVS.l1

      Estimate Std. Error t value Pr(>|t|)
rVCB.l1  0.03020   0.09863   0.306   0.760
rCTG.l1  0.04320   0.07934   0.545   0.586
rPVC.l1 -0.18977   0.07432  -2.553   0.011 *
rPVS.l1  0.09387   0.09488   0.989   0.323
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Residual standard error: 2.491 on 493 degrees of freedom
 Multiple R-Squared: 0.01751, Adjusted R-squared: 0.009539
 F-statistic: 2.197 on 4 and 493 DF, p-value: 0.0683

```
Estimation results for equation rPVS:
=====
rPVS = rVCB.l1 + rCTG.l1 + rPVC.l1 + rPVS.l1

      Estimate Std. Error t value Pr(>|t|)
rVCB.l1  0.01119   0.08015   0.140   0.8890
rCTG.l1  0.01072   0.06447   0.166   0.8680
rPVC.l1 -0.10352   0.06039  -1.714   0.0871 .
rPVS.l1 -0.02537   0.07710  -0.329   0.7423
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Residual standard error: 2.024 on 493 degrees of freedom
 Multiple R-Squared: 0.02117, Adjusted R-squared: 0.01323
 F-statistic: 2.666 on 4 and 493 DF, p-value: 0.03183

Covariance matrix of residuals:

| | | | | |
|--------|--------|--------|--------|--------|
| rVCB | rCTG | rPVC | rPVS | |
| 1.4469 | 0.6252 | 0.6102 | 0.5622 | |
| rCTG | 0.6252 | 2.5900 | 1.5771 | 1.4109 |
| rPVC | 0.6102 | 1.5771 | 6.2057 | 4.0666 |
| rPVS | 0.5622 | 1.4109 | 4.0666 | 4.0897 |

Correlation matrix of residuals:

| | | | | |
|--------|--------|--------|--------|--------|
| rVCB | rCTG | rPVC | rPVS | |
| 1.0000 | 0.3229 | 0.2036 | 0.2311 | |
| rCTG | 0.3229 | 1.0000 | 0.3934 | 0.4335 |
| rPVC | 0.2036 | 0.3934 | 1.0000 | 0.8072 |
| rPVS | 0.2311 | 0.4335 | 0.8072 | 1.0000 |