BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC NGOẠI THƯƠNG

ĐỀ TÀI

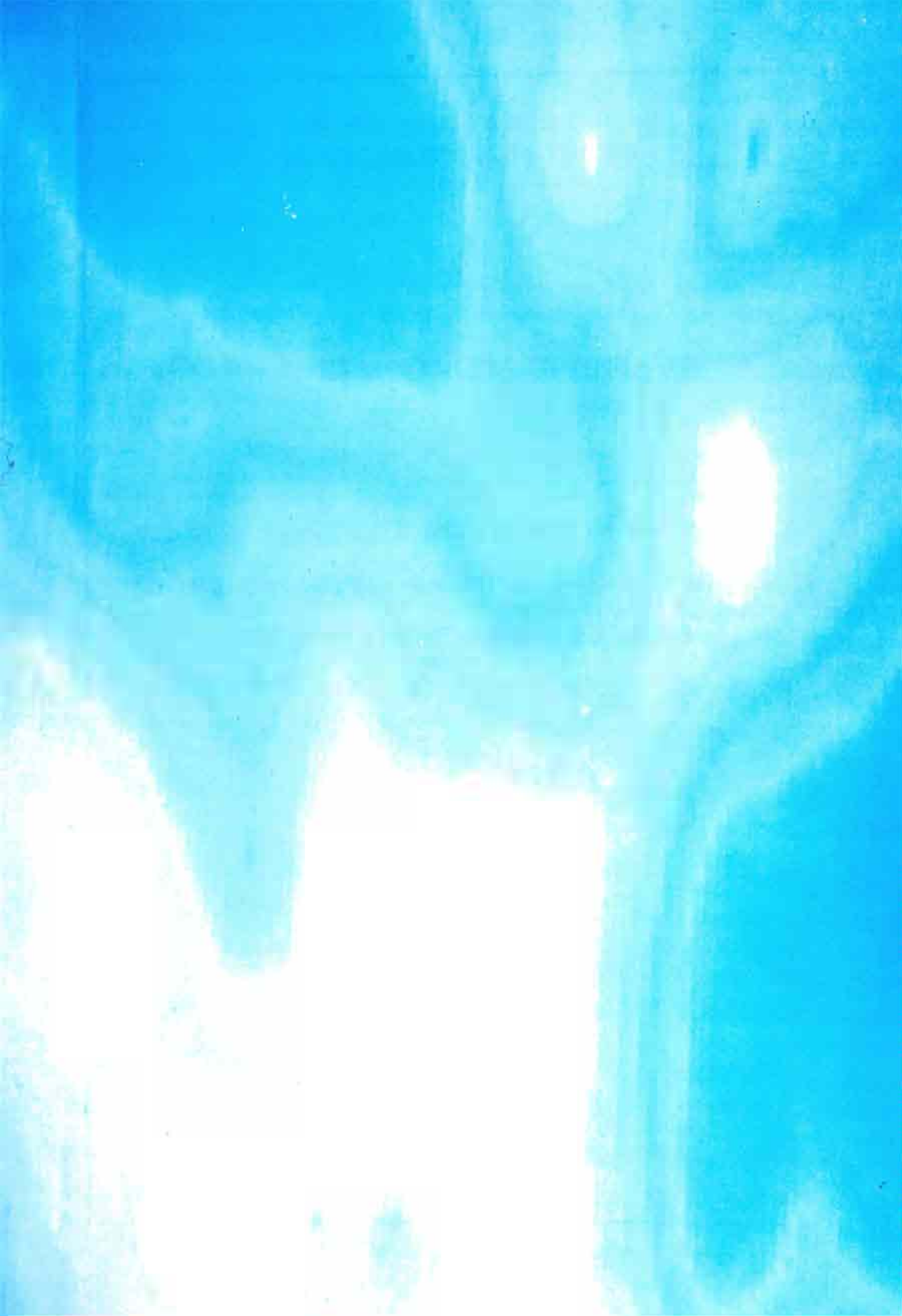
NGHIÊN CỨU KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG

MÔ HÌNH CHUỖI THỜI GIAN DÙNG ĐỂ DỰ BÁO BIẾN ĐỘNG GIÁ CHỨNG KHOÁN VÀ ÁP DỤNG VÀO THỊ TRƯỜNG CHỨNG KHOÁN VIỆT NAM

MÃ SỐ: NT2007-02

Chủ nhiệm đề tài: ThS PHÙNG DUY QUANG
Những người tham gia:
KS Nguyễn Tiến Thành
ThS Nguyễn Dương Nguyên
ThS Đinh Lê Hùng
ThS Lê Văn Tuấn

Hà nội, 2008



BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC NGOẠI THƯƠNG

------------- &&&-------------

ĐỀ TÀI NGHIÊN CỨU KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG

MÔ HÌNH CHUỖI THỜI GIAN DÙNG ĐỂ DỰ BÁO BIẾN ĐỘNG GIÁ CHỨNG KHOÁN VÀ ÁP DỤNG VÀO THỊ TRƯỜNG CHỨNG KHOÁN VIỆT NAM

MÃ SỐ: NT2007 - 02

Chủ nhiệm đề tài: **ThS Phùng Duy Quang**

Thành viên: KS Nguyễn Tiến Thành - *ĐH Bách Khoa Hà nội*
ThS Nguyễn Dương Nguyễn - *ĐH Ngoại Thương*
ThS Đinh Lê Hùng - *Ngân hàng Công Thương VN*
ThS Lê Văn Tuấn - *ĐH Thương Mại*

HÀ NỘI, 08. 2008

MỤC LỤC

LỜI NÓI ĐẦU

1. Tính cấp thiết của đề tài

Việt Nam đã và đang thiết lập nền kinh tế thị trường định hướng XHCN. Cơ chế quản lý kinh tế, tài chính đã và đang đổi mới sâu sắc, toàn diện với mục tiêu tăng trưởng với tốc độ cao, bền vững, xây dựng đất nước giàu mạnh. Chính sách kinh tế phải được hoạch định phù hợp với điều kiện cụ thể của Việt Nam. Để có được chính sách kinh tế năng động, hợp lý, có hiệu quả, dự báo kinh tế là một công cụ hữu ích làm cơ sở khoa học, có căn cứ để đưa ra các quyết định và xây dựng các giải pháp. Phân tích và dự báo có khả năng giúp cho việc nhận biết những vấn đề phát sinh khi vận hành chính sách thực tế. Việc áp dụng các phương pháp và mô hình dự báo còn giúp việc đánh giá và lựa chọn chính sách và giải pháp phù hợp. Sử dụng các mô hình dự báo kinh tế có thể nâng cao năng lực hoạch định chính sách của những người hoạch định và lựa chọn chính sách của các doanh nghiệp. Từ đó, các doanh nghiệp có thể linh hoạt điều chỉnh đối sách kinh doanh của mình phù hợp với điều kiện kinh tế xã hội.

Ngày nay, cùng với sự phát triển của khoa học và công nghệ, các phương pháp dự báo định lượng, đặc biệt là phương pháp sử dụng các quá trình ngẫu nhiên dùng để dự báo (trong đó có việc sử dụng các mô hình chuỗi thời gian) được sử dụng nhiều trong thực tiễn và thu được nhiều kết quả đáng tin cậy. Có khá nhiều kết quả dự báo có tính ổn định cao và giúp ích rất nhiều trong công tác hoạch định chính sách.

Thị trường chứng khoán Việt Nam tuy mới ra đời (năm 2000) và rất còn non trẻ so với các thị trường khác, nhưng với hơn 1500 phiên giao dịch bắt đầu từ phiên giao dịch đầu tiên 28 tháng 7 năm 2000 cũng đã cung cấp cho các nhà nghiên cứu một khối lượng lớn số liệu đủ để phân tích, ước lượng và dự báo. Để thấy được tầm quan trọng của các mô hình ARCH và GARCH trong việc giải thích được các đặc trưng của biến động trên các chuỗi số liệu chứng khoán Việt Nam, đòi hỏi nhiều thời gian và công sức của các nhà nghiên cứu về mặt lý thuyết cũng như mặt ứng dụng. Liên quan tới lĩnh vực này là bài toán dự báo biến động về giá của các chứng khoán giao dịch trên thị trường cũng cần được đầu tư vào nghiên cứu. Trong hiện trạng như vậy, việc tìm tòi và nghiên cứu, triển khai ứng dụng các lý thuyết về các mô hình chuỗi thời gian vào các chuỗi thời gian tài chính Việt Nam là rất thiết thực và hiệu quả. Từ đó, có thể tìm kiếm các công cụ định lượng giúp cho các nhà đầu tư có cơ sở và căn cứ khoa học để tiến hành các hoạt động đầu tư của mình vào thị trường chứng khoán Việt Nam.

Với những lý do trên, nhóm tác giả lựa chọn đề tài " **Mô hình chuỗi thời gian dùng để dự báo biến động giá chứng khoán và áp dụng vào thị trường chứng khoán Việt Nam** " để góp phần giải quyết những nhu cầu bức xúc của doanh nghiệp, cũng như hoạt động quản lý kinh tế, quản lý kinh doanh về dự báo thị trường ...

2. Tình hình nghiên cứu trong và ngoài nước

2.1.Tổng quan tình hình nghiên cứu thuộc lĩnh vực của đề tài

Lý thuyết về các quá trình ngẫu nhiên, đặc biệt là cơ sở lý thuyết về chuỗi thời gian đã được nghiên cứu sâu ở nhiều nước tiến tiến trên thế giới. Sử dụng các kết quả lý thuyết về các mô hình chuỗi thời gian AR, ARMA cũng đã được triển khai nhiều trong khoa học dự báo và thu được nhiều kết quả đáng tin cậy (dạng sách và các ấn phẩm). Kể từ khi mô hình ARCH do Engle phát hiện ra năm 1982, các kết quả lý thuyết đặc sắc của các mô hình kiểu ARCH và GARCH đã được triển khai và ứng dụng trong nhiều lĩnh vực, đặc biệt là lĩnh vực tài chính.

Ở nước ta, các kết quả nghiên cứu về chuỗi thời gian còn riêng lẻ trong việc sử dụng các mô hình chuỗi thời gian trong dự báo một số đối tượng như dự báo về giá, về nhu cầu, ... Tuy nhiên, các mô hình chuỗi thời gian chỉ dừng lại ở dạng tuyến tính AR, ARMA, chưa có nhiều kết quả ứng dụng đối với các mô hình phi tuyến kiểu ARCH và GARCH. Gần đây đã có một số kết quả phân tích về hiệu ứng ARCH và GARCH trên các thời gian chứng khoán của Việt Nam những cũng chưa áp dụng kết quả nhận dạng vào bài toán thực tiễn.

Trước thực trạng như vậy, vấn đề phân tích, xây dựng các mô hình chuỗi thời gian ARCH và GARCH và ứng dụng trong bài toán dự báo - hiện đang là vấn đề thời sự trong lĩnh vực Toán ứng dụng, cần thiết được đặt ra và giải quyết kịp thời, từ đó ứng dụng các kết quả vào các bài toán thực tiễn.

2.2. Danh mục các công trình liên quan

a. Của chủ nhiệm và những người tham gia thực hiện đề tài

1. *Phùng Duy Quang*, "Các mô hình chuỗi thời gian ARMA và ARCH. Ứng dụng vào bài toán dự báo", đề tài Thạc sỹ Toán học, chuyên ngành Lý thuyết xác suất và Thống kê toán, ĐHBK Hà nội, 2006.

2. *Nguyễn Hồ Quỳnh, Phùng Duy Quang, Nguyễn Tiến Thành,*" Mô hình chuỗi thời gian ARCH và ứng dụng dự báo biến động chuỗi thời gian tài chính ", *Kỷ yếu Hội nghị NCKH kỷ niệm 50 năm thành lập trường ĐHBK Hà nội*, 2006.

3. *Phùng Duy Quang*, "Dự báo ngắn hạn giá hàng xuất khẩu bằng phương pháp làm trơn Holt – Winters ", *Tạp chí Kinh tế Đối ngoại*, trang 46, số 23/2007.

b. Của Những người khác

1. *Vương Quân Hoàng* (2004), "Hiệu ứng GARCH trên dãy lợi suất thị trường chứng khoán Việt Nam 2000 – 2003 ", *Tạp chí ứng dụng Toán học, tập II, số 1,* trang 15 -30.

2. *Vũ Hoài Chương* (2002), " Chuỗi thời gian và các kết luận thống kê", Đề tài NCKH, Viện CNTT – Viện KH & CN Việt Nam.

3. *Chong C. W., Ahmad M. I. and Abdullah M. Y.*(1999), "Performance of GARCH Models in Forecasting Stock Market Volatility ", *Journal of forecating,* 18, pp. 333 – 343.

4. *Jorion P.* (1995), "Predicting volatility in the foreign exchange market", *Journal of Finance, 50,* pp. 507 – 528.

5. *Najand M.* (2002), "Forecasting Stock Index Futures Price Volatility: Linear vs. Nonlinear Models ", *The financial Review*, 37, pp. 93 -104.

6. *Poon S. , Granger C. J. W.* (2003), "Forecasting volatility in financial markets: A review ", *Journal of Economic Liturature, 41,* pp.478 – 639.

3. Mục tiêu đề tài

1. Tổng quan về các phương pháp dự báo, đặc biệt là dự báo theo mô hình chuỗi thời gian.

2. Nghiên cứu thực trạng biến động giá chứng khoán của thị trường chứng khoán Việt Nam bằng các mô hình ngẫu nhiên và mô hình chuỗi thời gian.

3. Thiết kế phần mềm chuyên dụng phục vụ phân tích các hoạt động chứng khoán dựa trên các mô hình chuỗi thời gian phục vụ cho giảng dạy tại Đại học Ngoại Thương và các môn Tài chính và Chứng khoán.

4. Cách tiếp cận, phương pháp nghiên cứu, phạm vi nghiên cứu

* Cách tiếp cận, phương pháp nghiên cứu:

- Phương pháp tiếp cận hệ thống; phương pháp Xác suất và thống kê Toán.
- Phương pháp tổng hợp và phân tích định lượng.
- Phương pháp tiên đề và phương pháp mô phỏng; phương pháp đối chiếu - so sánh.

* Đối tượng, phạm vi nghiên cứu:

- Đề tài chỉ tập trung nghiên cứu phương pháp định lượng cơ bản là phương pháp dùng các mô hình chuỗi thời gian để dự báo.
- Đối tượng nghiên cưu là giá của chứng khoán giao dịch trên một số thị trường chứng khoán trên thế giới và Việt Nam
- Các mô hình chuỗi thời gian áp dụng cho dự báo biến động giá chứng khoán phục vụ cho công tác giảng dạy và học tập tại Đại học Ngoại Thương.

5. Kết cấu của đề tài

Ngoài phần mở đầu và kết luận, danh mục tài liệu tham khảo, nội dung của đề tài dàn trải thành 3 phần sau:

Chương 1. Cơ sở khoa học của dự báo

Chương 2. Các mô hình chuỗi thời gian tài chính dạng ARCH

Chương 3. Dự báo biến động giá chứng khoán và áp dụng vào thị trường chứng khoán Việt Nam

CHƯƠNG 1. CƠ SỞ KHOA HỌC CỦA DỰ BÁO

Chương đầu tiên của đề tài, tập trung nghiên cứu cơ sở lý luận cũng như cơ sở toán học của khoa học dự báo. Nội dung chính của chương gồm các phần: Giới thiệu tổng quan về dự báo; các bước cần thực hiện trong quá trình dự báo; phân loại các kiểu dự báo; độ chính xác của dự báo; phương pháp bình phương cực tiểu; kỳ vọng có điều kiện và không gian $L^2(\Omega, F, P)$.

1.1. Tổng quan về dự báo

1.1.1. Tại sao phải dự báo

Việc dự báo các đại lượng biến thiên đóng vai trò rất quan trọng trong khoa học và kỹ thuật. Chúng giúp cho những người ra quyết định, những nhà doanh nghiệp tiên đoán một cách khoa học xu hướng phát triển trong tương lai của các đại lượng đó và từ đó ta có thể lên kế hoạch – hoạch định chính sách, phương hướng đầu tư – phát triển đúng đắn.

Trong lĩnh vực quản lý và các tình huống thuộc về quản lý của các doanh nghiệp rất cần thiết lên kế hoạch sản xuất từng tuần, từng tháng và từng năm. Đây là một trong những vấn đề quan trọng đóng vai trò quyết định thành công của doanh nghiệp. Vì vậy để đảm bảo cho việc sử dụng hiệu quả các thiết bị và cơ sở vật chất, sử dụng nguồn vốn vào đầu tư mở rộng sản xuất, thuê mướn nhân công, mở rộng sản xuất hàng hoá nào, hạn chế sản xuất hàng hoá nào, đòi hỏi chúng ta cần dự báo các đối tượng ảnh hưởng đến các mặt đó với mức độ càng chính xác càng tốt.

Bài toán dự báo cũng đặc biệt quan trọng ở ngành bưu chính viễn thông. Đó là ngành dịch vụ với qui mô lớn, sử dụng các thiết bị đắt tiền, đòi hỏi việc đầu tư về cơ sở hạ tầng rất lớn và liên tục. Đặc biệt là thiết bị hiện đại phục vụ cho trang bị hệ thống thông tin liên lạc. Ngày này, các doanh nghiệp kinh doanh trong lĩnh vực bưu chính viễn thông đang cạnh tranh khốc liệt nhằm giành giật từng khách hàng sử dụng các dịch vụ do họ cung cấp. Tính huống này đặt ra đòi hỏi các doanh nghiệp phải không ngừng trang bị các thiết bị hiện đại, đồng thời giảm giá của các dịch vụ xuống, nhưng họ vẫn phải làm thế nào để mình kinh doanh vẫn có lãi. Chính vì điều này, cần thiết dự báo nhu cầu với mức độ càng chính xác bao nhiêu càng tốt.

Trong lĩnh vực khí tượng – thuỷ văn, việc dự báo thời tiết như nhiệt độ, nắng mưa, lũ lụt, ... sẽ giúp ích nhiều cho nền kinh tế cũng như tránh được những thiệt hại to lớn do thiên nhiên gây ra.

Trong lĩnh vực tài chính – tiền tệ, nếu ai biết được xu hướng tăng giảm tỷ giá của 1 loại tiền tệ hay giá của một cổ phiếu, sự biến động lên xuống của giá cổ phiếu, tỷ giá của đồng Việt nam/ USD chắc chắn mang lại nhiều lợi ích cho người đó.

Một lĩnh vực không kém quan trọng trong nền kinh tế quốc dân nếu biết được nhu cầu sử dụng các loại phương tiện giao thông như ô tô, xe máy, xe đạp, ... cùng với số lượng bao nhiêu sẽ giúp ích cho các nhà quản lý có kế hoạch phát triển mạng lưới giao thông của một thành phố, không những khắc phục được tình trạng ùn tắc giao thông hiện nay mà còn tăng hiệu quả kinh tế cho nền kinh tế.

Phương hướng chung là có thể dự báo chính xác hơn một số lượng các sự kiện, đặc biệt là các sự kiện trong môi trường kinh doanh sẽ cung cấp cơ sở tốt hơn cho việc lên kế hoạch. Từ các phân tích trên ta cần nhấn mạnh 2 điểm quan trọng trong công tác dự báo: *Thứ nhất*, các dự báo không chỉ thành công trong một lĩnh vực nhất định mà nó có thể áp dụng trực tiếp cho nhiều lĩnh vực của đời sống xã hội. *Thứ hai*, trong công tác dự báo cần phân biệt được hai lớp nhân tố ảnh hưởng đến một đối tượng nhất định: các nhân tố bên ngoài không điều khiển được và các nhân tố điều khiển được. Sự thành công của bất kỳ một cơ quan hay một doanh nghiệp nào cũng phụ thuộc vào hai lớp nhân tố đó. Các phương pháp dự báo được ứng dụng để xem xét các ảnh hưởng của lớp nhân tố thứ nhất lên đối tượng nghiên cứu. Trong khi ra quyết định sử dụng các lớp nhân tố thứ 2. Lên kế hoạch là sự kết hợp hai lớp nhân tố này. Điểm quan trọng trong công tác dự báo là xác định được nhân tố nào là trọng yếu, nhân tố nào là thứ yếu; từ đó áp dụng phương pháp dự báo thích hợp. Do vậy, dự báo cũng là một phần quan trọng trong công tác lên kế hoạch của một cơ quan hay doanh nghiệp.

1.1.2. Tổng quan về kỹ thuật dự báo

Các tình huống dự báo thay đổi về phạm vi thời gian, không gian, các đối tượng, dữ liệu và các khía cạnh khác. Để ứng dụng các phương pháp khác nhau, nhiều kỹ thuật dự báo sẽ được áp dụng, các kỹ thuật này phân thành 2 nhóm phương pháp: phương pháp định lượng và phương pháp định tính.

* *Phương pháp định lượng:* Nhóm phương pháp này được sử dụng khi các thông tin định lượng là có giá trị. Đối với nhóm phương pháp này ta thường sử dụng 2 mô hình: mô hình chuỗi thời gian và mô hình giải thích.

- Đối với mô hình chuỗi thời gian: Dự báo là quá trình ngoại suy dữ liệu từ quá khứ vào tương lai.

- Đối với mô hình giải thích: Sử dụng mối liên hệ giữa biến giải thích và biến độc lập thông qua các phương trình, một trong lớp phương trình đó là các phương trình hồi quy. Các thông tin đều cho trên các mẫu dữ liệu của các biến đó.

* *Phương pháp định tính:* Nhóm phương pháp này được sử dụng khi các thông tin định lượng ít hoặc không có giá trị nhưng các thông tin định tính lại có giá trị. Chẳng hạn xét các ví dụ:

- Dự báo sự lan truyền thông tin liên lạc năm 2020.
- Dự báo sự tăng giá dầu ảnh hưởng như thế nào đến thị hiếu tiêu thụ dầu trên thế giới, ...

Ngoài ra, trên thực tế ta còn gặp các tình huống không dự báo được, tình huống này xảy ra khi các thông tin định tính hoặc định lượng ít hoặc không có giá trị. Chẳng hạn:

- Dự báo tác động của du lịch liên hành tinh tới nền kinh tế quốc dân.
- Dự báo khám phá ra cái mới như 1 nguồn năng lượng mới,...

Các phương pháp định lượng có thể được ứng dụng khi 3 điều kiện sau đây được thoả mãn:

1. Thông tin về quá khứ là có giá trị.
2. Các thông tin đó có thể lượng hoá dưới dạng dữ liệu số.
3. Có thể giả thiết một vài ảnh hưởng của quá khứ còn tiến triển vào tương lai.

Điều kiện thứ ba được hiểu là giả thuyết liên tục của một sự kiện, nó là giả thuyết cơ bản của tất cả các phương pháp định lượng và nhiều phương pháp dự báo định tính khác. Ngày nay, các phương pháp định lượng được ứng dụng nhiều trong các lĩnh vực khác nhau, cho các mục đích khác nhau, cho các đối tượng khác nhau. Trong mỗi trường hợp đó mỗi phương pháp có độ chính xác và giá trị khác nhau. đề tài chỉ tập trung nghiên cứu các phương pháp định lượng dùng mô hình chuỗi thời gian để dự báo.

1.2. Các bước cần thực hiện trong quá trình dự báo

Quy trình dự báo chuỗi thời gian bao gồm các bước sau:

Bước 1. *Xác định mục tiêu của dự báo.* Đây là một khía cạnh khó của bài toán dự báo; nó bao gồm cách hiểu sâu sắc về ứng dụng của các dự báo, đối tượng nào được dự báo, khu vực và phạm vi được dự báo. Khía cạnh này chính là vấn đề đặt bài toán, việc xác định bài toán đúng giúp chúng ta tìm ra được lời giải tối ưu. Do vậy, trong thực tiễn người làm công tác dự báo phải xác định rõ mục tiêu cụ thể của các dự báo, từ đó lên kế hoạch làm các dự báo. Ba mục tiêu chính của dự báo cần đạt được là:

* Đối tượng dự báo: giá tôm xuất khẩu vào thị trường Mỹ, tỷ giá của đồng Việt Nam với đô la Mỹ, tỷ giá cổ phiếu, biến động giá cổ phiếu, ...

* Khu vực dự báo: Tuỳ theo lĩnh vực, ngành, hay một đơn vị nào đó.

* Thời gian dự báo: 1 ngày, 1 tuần, 1 tháng hay là 1 năm, ...

Ngoài 3 mục tiêu trên, để sử dụng phương pháp dự báo thích hợp, cần phân tích đầy đủ các yếu tố ảnh hưởng đến đối tượng cần dự báo: Các nhân tố không điều khiển được và các nhân tố điều khiển được. Việc phân tích 2 lớp nhân tố này một cách đầy đủ giúp ta chọn lựa đối tượng dự báo phù hợp, từ đó lựa chọn mô hình phù hợp.

Bước 2. *Thu thập và phân loại dữ liệu.* Sau khi xác định xong đối tượng dự báo, cần thu thập dữ liệu theo 2 loại:

- Dữ liệu thống kê.
- Tri thức kinh nghiệm và các tri thức của các chuyên gia. Các tri thức này chủ yếu sử dụng trong dự báo định tính.

Quá trình thu thập dữ liệu thống kê cần xác định được các yếu tố sau:

- Dữ liệu dạng chéo (không phụ thuộc thời gian).
- Dữ liệu dạng chuỗi thời gian.
- Kích thước mẫu dữ liệu.
- Đối với dữ liệu chuỗi thời gian, cần xác định thời điểm bắt đầu và thời điểm kết thúc.

Hai yếu tố đầu quyết định việc chọn loại mô hình cho chuỗi dữ liệu, còn 2 yếu tố sau việc chọn mô hình phù hợp.

Bước 3. *Phân tích thô số liệu.*

Mục đích của phân tích thô là từ dữ liệu cung cấp những thông tin gì về đối tượng dự báo. Trước hết, vẽ đồ thị dữ liệu để có cái nhìn trực giác về đối tượng nghiên cứu. Đối với chuỗi thời gian ngoài đồ thị chuỗi số liệu, ta còn phác thảo đồ thị hàm tự tương quan và tự tương quan riêng mẫu; các kết quả quan sát được từ các loại đồ thị này cho ta kết luận về tính dừng của chuỗi thời gian (giả thiết mang tính bản chất của các mô hình) mà chúng ta sẽ giới thiệu ở mục 2.1.

Các thống kê đơn giản cho mẫu dữ liệu cũng được tính toán như trung bình, phương sai, độ lệch chuẩn, giá trị lớn nhất, giá trị nhỏ nhất, ...

Quá trình phân tích thô một chuỗi thời gian nhằm phát hiện các thành phần hợp thành nên chuỗi thời gian, các thành phần đó có thể là:

- Khuynh hướng phát triển của chuỗi thời gian, sự tăng lên hay giảm xuống của chuỗi thời gian.
- Chuỗi thời gian có chứa yếu tố mùa hay không ?
- Chuỗi thời gian có chứa yếu tố tuần hoàn hay không ?
- Thành phần bất quy tắc.

Ngoài ra, các điểm bất thường của chuỗi thời gian cũng được xem xét và được giải thích 1 cách đầy đủ.

Do đó, nội dung của sự phân tích thô là nhằm phát hiện các thành phần của chuỗi thời gian. Từ đó nhận dạng các thành phần đó theo các nguyên tắc hợp thành của các thành phần của chuỗi thời gian, sau đó mới chọn mô hình phù hợp cho dữ liệu.

Bước 4. *Xác định kỹ thuật dự báo.*

Các phương pháp dự báo có thể tạm phân làm 3 loại:

* Ngoại suy chuỗi thời gian
* Phân tích hồi quy
* Các phương pháp khác

Phương pháp dự báo thường được chọn tương ứng với đặc điểm của đối tượng cần dự báo và các yếu tố liên quan, ứng với các dữ liệu thu được. Khi chọn phương pháp dự báo căn cứ vào việc lựa chọn mô hình phù hợp với dữ liệu. Để thu được các giá trị dự báo với độ tin cậy cao, điều quan trọng là phải chọn lựa mô hình phù hợp với dữ liệu. Muốn vậy, căn cứ vào các kết qủa phân tích thô để ta lựa chọn một mô hình phù hợp, có tính khả thi cao.

Sau khi lựa chọn mô hình, các tham số của mô hình cũng được ước lượng trên cơ sở dữ liệu. Trong các phương pháp ước lượng lựa chọn phương pháp nào hiệu qủa nhất.

Bước 5. *Kiểm định sự phù hợp của mô hình.* Việc kiểm định sự phù hợp của mô hình nhằm trả lời 2 câu hỏi:

* Mô hình đã phù hợp rồi, thì sử dụng vào đâu, nhằm mục đích gì?

* Mô hình chưa phù hợp với dữ liệu, thì chúng ta có thể nhận dạng lại hay không?

Trong trường hợp mô hình đã phù hợp với dữ liệu, chuyển sang bước 6. Còn trong trường hợp ngược lại ta lặp lại bước 3, bước 4, bước 5 cho đến khi xây dựng được mô hình phù hợp thì chuyển sang bước 6.

Bước 6. *Xác định các giá trị dự báo theo mô hình.*

Sau khi đã chọn được các mô hình phù hợp với dữ liệu, sử dụng mô hình để tính toán các giá trị dự báo tương lai của dữ liệu. Khi đó phân tích các ưu, nhược điểm của từng mô hình. Đánh giá độ chính xác của phương pháp dự báo thông qua các thống kê đã có. Từ đó chọn lựa mô hình tối ưu nhất, tính toán các giá trị dự báo theo mô hình đó.

1.3. Phân loại các kiểu dự báo

1.3.1. Phân loại theo quy mô vùng dự báo

Việc phân loại này mang tính chất tương đối, nhằm mục đích sử dụng các công cụ định lượng hoặc định tính trong quá trình dự báo. Chẳng hạn, dự báo tổng kim ngạch xuất khẩu của cả nước là dự báo ở cấp vĩ mô. Trong khi dự báo giá tôm xuất khẩu vào thị trường Nhật Bản là dự báo ở cấp vi mô.

1.3.1.1. Dự báo cấp vi mô

Dự báo cấp vi mô là dự báo cho một vùng hay cho một khu vực nhỏ. Có thể nói là dự báo cho một ngành cụ thể.

Thí dụ: xây dựng chiến lược kinh doanh, lắp đặt các trang thiết bị, xây dựng kế hoạch xuất khẩu hàng hoá, ... đều có thể sử dụng dự báo ở cấp vi mô.

1.3.1.2. Dự báo cấp vĩ mô

Có nhiều phương pháp thống kê sử dụng ở cấp vĩ mô. Có phương pháp định tính cũng như phương pháp định lượng. Dự báo ở cấp vĩ mô là sự phối kết hợp nhiều phương pháp, trong đó có sử dụng các kết quả của dự báo ở cấp vi mô. Do vậy, cần có sự nghiên cứu đầy đủ, chi tiết và một cách toàn diện đối tượng dự báo. Chẳng hạn như dự báo tổng thu nhập quốc dân, tốc độ phát triển kinh tế của đất nước, ...

Nhìn chung, việc phân loại dự báo theo qui mô, giúp cho các nhà quản lý sử dụng các mô hình toán học thích hợp vào các ngành, lĩnh vực cụ thể. Thông thường các kết quả của dự báo ở cấp vi mô được sử dụng trong dự báo ở cấp vĩ mô, nhưng cũng có nhân tố được sử dụng để dự báo ở cấp vi mô nhưng không sử dụng để dự báo ở cấp vĩ mô. Do vậy, phân loại theo khu vực cũng giúp cho việc chọn lựa mô hình phù hợp.

1.3.2. Phân loại theo thời kỳ dự báo

Sử dụng các mô hình toán học vào công tác dự báo thường căn cứ nhiều vào thời kỳ dự báo. Thời kỳ dự báo của một đối tượng (như nhu cầu, đơn giá, khối lượng của một loại hàng hoá ...) được phân thành dự báo ngắn hạn, dự báo trung hạn và dự báo dài hạn. Sự phân loại này tuỳ thuộc và độ dài của thời kỳ dự báo được sử dụng và mục đích dự báo.

1.3.2.1. Dự báo ngắn hạn

Dự báo này là dự báo cho một hoặc hai thời kỳ tiếp theo (ngày, tháng, ...)

Dự báo ngắn hạn đòi hỏi thông tin tương đối chính xác, có xét tới các nhân tố bên ngoài ảnh hưởng tới đối tượng dự báo. Chẳng hạn dự báo giá tôm xuất khẩu vào thị trường Mỹ của tháng 12/ 2006.

1.3.2.2. Dự báo trung hạn

Các dự báo được tiến hành từ 3 đến 5 thời kỳ. Các phương pháp dự báo chuỗi thời gian thường sử dụng cho dự báo ngắn hạn và trung hạn. Triết lý của dự báo theo mô hình chuỗi thời gian là dùng thông tin của quá khứ thể hiện trong tập quan sát $\mathscr{X} = \{x_t; t = \overline{1,n}\}$- sau này ta sẽ gọi là *chuỗi thời gian,* cho đến hiện tại để suy đoán các giá trị tương lai.

1.3.2.3. Dự báo dài hạn

Dự báo dài hạn cho khoảng 5 thời kỳ trở lên. Nó liên quan tới kế hoạch và chiến lược ở tầm vĩ mô. Trong trường hợp này, việc dự báo bằng cách sử dụng mô hình chuỗi thời gian không còn phù hợp nữa. Dự báo này về cơ bản được thực hiện bằng cách dự đoán gián tiếp sử dụng các mô hình mà có thể giải thích được mối quan hệ giả định của các yếu tố khác đối với đối tượng dự báo. Trong dự báo dài hạn người ta thường sử dụng các phương trình hồi quy.

Việc lựa chọn các mô hình dự báo người ta căn cứ vào thời kỳ dự báo và mục đích của dự báo. Không thể áp dụng 1 phương pháp dự báo cho tất cả các thời kỳ ngắn hạn, trung hạn và dài hạn. Do vậy, trong thực hành sử dụng các phương pháp dự báo cần chú ý :

- Đối tượng cần dự báo theo bao nhiêu thời kỳ.
- Có thể sử dụng nhiều phương pháp dự báo cho từng thời kỳ.
- Cần điều chỉnh các dự báo khi sử dụng các phương pháp dự báo theo 2 thời kỳ.

1.4. Độ chính xác của dự báo

Khi thực hiện dự báo theo 1 chuỗi thời gian $\mathcal{X}$ cần thiết phải trả lời câu hỏi: liệu mô hình đó có phù hợp với dữ liệu không, tiêu chuẩn nào để đánh giá sự phù hợp đó, độ chính xác của dự báo được xác định như thế nào. Để trả lời câu hỏi đó, ta sử dụng các thống kê để đánh giá độ chính xác của dự báo thông qua sai số của giá trị dự báo và giá trị thực tế của chuỗi thời gian. Độ chính xác dự báo được xem là tiêu chuẩn để chọn lựa mô hình tối ưu nhất.

1.4.1. Các thống kê đo độ chính xác của dự báo

Giả sử X_t là quan sát thực tế tại thời điểm t, F_t là giá trị dự báo của X_t tại thời điểm t trên cơ sở tập thông tin $\mathcal{X}_t = \{x_j; j = \overline{1,t}\}$. Khi đó sai số dự báo bước 1 là

$$e_t = X_t - F_t \qquad (1.1)$$

Giả sử có n giá trị thực tế $X_1, X_2, \ldots, X_n$ và n giá trị dự báo $F_1, F_2, \ldots, F_n$ tương ứng thì chúng ta sẽ có n giá trị sai số dự báo bước 1: $e_1, e_2, \ldots, e_n$.

Thông thường, để đánh giá độ chính xác của 1 phương pháp dự báo, ta thường sử dụng các thống kê sau:

* Sai số trung bình (mean error):

$$ME = \frac{1}{n}\sum_{t=1}^{n} e_t \tag{1.2}$$

* Trung bình trị tuyệt đối sai số (mean absolute error):

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |e_t| \tag{1.3}$$

* Trung bình bình phương sai số (mean square error):

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2 \tag{1.4}$$

Khi sử dụng các thống kê (1.2), (1.3) và (1.4) cần chú ý:

- Thống kê ME có thể nhỏ tuỳ ý. Mặc dầu các sai số có thể lớn về giá trị tuyệt đối (sai số dương và sai số âm triệt tiêu nhau). Do vậy thống kê ME chưa đặc trưng cho độ chính xác của dự báo. Thống kê MAE và MSE khắc phục được điều này bằng cách lấy trung bình trị tuyệt đối sai số và trung bình bình phương sai số.

- Thống kê MAE thuận lợi là dễ sử dụng và giải thích ý nghĩa độ chính xác của dự báo đối với những người không có nhiều chuyên môn về toán học.

- Thống kê MSE dễ dàng nghiên cứu về mặt toán học.

Tuy nhiên các thống kê trên đều phụ thuộc vào đơn vị đo dữ liệu, nên khó có thể so sánh cùng 1 phương pháp dự báo nhưng áp dụng cho 2 loại đối tượng khác nhau thì phương pháp nào hay hơn. Khắc phục nhược điểm này, ta sử dụng sai số tương đối theo tỷ lệ phần trăm:

* Sai số tương đối (Relation error) tại thời điểm t :

$$PE_t = \frac{X_t - F_t}{X_t}.100\%$$

* Trung bình các sai số tương đối:

$$MPE = \frac{1}{n}\sum_{t=1}^{n} PE_t$$

* Trung bình trị tuyệt đối sai số tương đối:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} |PE_t|$$

Tương tự như phân tích trên MAPE khắc phục được nhược điểm của MPE là giá trị có thể nhỏ tuỳ ý nhưng các sai số tương đối có thể lớn về trị tuyệt đối. Trong thực hành thường sử dụng 2 thống kê MSE và MAPE để đánh giá độ chính xác của phương pháp dự báo.

1.4.2. Phương pháp đánh giá độ chính xác của dự báo ngoài mẫu

Phương pháp này chia dữ liệu thành 2 phần:

- Tập dữ liệu dùng để xây dựng mô hình (tập dữ liệu gốc)
- Tập dữ liệu kiểm tra.

Việc nhận dạng mô hình và ước lượng các tham số của mô hình sử dụng tập dữ liệu dùng để xây dựng mô hình. Dự báo được tiến hành trên tập dữ liệu kiểm tra. Từ chỗ tập dữ liệu kiểm tra không tham gia vào việc xây dựng mô hình, nên những giá trị dự báo thu được là sát thực mà không hề sử dụng các quan sát tại thời điểm dự báo. Độ chính xác của dự báo được tính toán sử dụng các thống kê trên cho các sai số dự báo của tập kiểm tra.

1.4.3. Khoảng dự báo cho dự báo điểm

Ngoài việc cung cấp giá trị dự báo cho chuỗi thời gian $\mathscr{X}$ – gọi là *dự báo điểm.* Các giá trị dự báo điểm này không phải hoàn toàn là giá trị thực tế của chuỗi thời gian. Do vậy, sử dụng khoảng dự báo để cung cấp cho người dùng các trường hợp xấu nhất và tốt nhất của các dự báo có thể có, tức là xác định khoảng biến thiên của các dự báo điểm – khoảng đó được gọi là *khoảng dự báo.* Xác định khoảng dự báo cho các dự báo điểm đều căn cứ vào phân bố xác suất của sai số dự báo. Bởi vì khoảng dự báo thường dựa trên thống kê MSE, nó cung cấp 1 ước lượng cho phương sai của sai số dự báo 1 bước.

Với giả thiết các sai số dự báo có phân bố chuẩn với trung bình 0, thì khoảng dự báo cho các dự báo điểm bước 1 là:

$$F_{n+1} \pm z\sqrt{MSE}$$

Trong đó F_{n+1} chính là giá trị dự báo của X_{n+1} trên cơ sở $\mathscr{X}$; z xác định độ rộng và xác suất của khoảng dự báo. Chẳng hạn với độ tin cậy 95 % thì z = 1,96 – tương ứng với khoảng dự báo chứa giá trị đúng với xác suất 95% hoặc với độ tin cậy 75 % thì z =1,15 – tương ứng với khoảng dự báo chứa giá trị đúng với xác suất 75%.

Trong trường hợp dự báo bước h, thay cho MSE ta sử dụng thống kê

$$MSE_h = \frac{1}{n-h}\sum_{t=h+1}^{n}\left(e_t^{(h)}\right)^2$$

Với $e_t^{(h)}$ là sai số dự báo bước h của giá trị dự báo h bước với giá trị thực tế tại thời điểm t. Khi đó khoảng dự báo cho dự báo bước h là

$$F_{n+h} \pm z\sqrt{MSE_h}$$

Trong đó F_{n+h} chính là giá trị dự báo bước h của X_n trên cơ sở $\mathscr{X}$.

1.5. Phương pháp bình phương cực tiểu

1.5.1. Bài toán hồi quy bội tuyến tính

Giả sử $X_1, X_2, \dots, X_k$ là k biến độc lập dùng để dự báo và Y là biến phụ thuộc cần dự báo. Sự phụ thuộc của Y vào các X_i $(i = \overline{1,k})$ thường rất phức tạp, nhưng trong các lĩnh vực kinh tế, nhiều sự phụ thuộc có thể chuyển về dạng tuyến tính. Một trong các mô hình tuyến tính được sử dụng nhiều trong kinh tế là mô hình hồi quy bội tuyến tính. Mô hình hồi quy bội tuyến tính khẳng định rằng Y phụ thuộc tuyến tính vào các X_i $(i = \overline{1,k})$ và sai số ngẫu nhiên ε theo biểu thức:

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + \varepsilon \qquad (1.5)$$

Trong đó β_i; $i = \overline{0,k}$ là các hệ số chưa biết.

Bài toán đặt ra là ước lượng các tham số β_i; $i = \overline{0,k}$ theo mô hình (1.5) dựa trên thông tin quan sát được của k + 1 biến ngẫu nhiên $X_1, X_2, \dots, X_k$ và Y. Giả sử thông tin quan sát là n quan sát độc lập đồng thời của k + 1 biến đó, các số liệu tuân theo mô hình sau:

$$y_t = \beta_0 + \beta_1 x_{t1} + \dots + \beta_k x_{tk} + \varepsilon_t,\ t = \overline{1,n};$$

hay

$$y_t = x_t^T\beta + \varepsilon_t,\ t = \overline{1,n}; \qquad (1.6)$$

dạng ma trận:

$$y = X.\beta + \varepsilon; \qquad (1.7)$$

trong đó

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix}; \; y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}; \beta = \begin{bmatrix} \beta_0 \\ \beta_2 \\ \dots \\ \beta_k \end{bmatrix}; \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{bmatrix};$$

$$x_t = \begin{bmatrix} 1 \\ x_{t1} \\ \dots \\ x_{tk} \end{bmatrix}; t = \overline{1, n}.$$

Trước hết, giả sử các sai số ngẫu nhiên thoả mãn điều kiện:

i) $E(\varepsilon_t) = 0$ (1.8)

ii) $Var(\varepsilon_t) = \sigma^2$ (1.9)

iii) $Cov(\varepsilon_i, \varepsilon_j) = 0$ với $\forall\ i \neq j$. (1.10)

điều kiện (1.8) và (1.9) tương đương với

$$Cov(\varepsilon) = E(\varepsilon\varepsilon^T) = \sigma^2 I_n \text{ (I_n là ma trận đơn vị cấp n)} \quad (1.11)$$

Trong trường hợp tổng quát (1.11) thay bởi $Cov(\varepsilon) = E(\varepsilon\varepsilon^T) = \Omega$, Ω là ma trận xác định dương, ta có *bài toán hồi quy tuyến tính tổng quát* [35]. Đề tài sử dụng các kết quả bài toán hồi quy bội tuyến tính (hồi quy tuyến tính cổ điển).

*** Phương pháp bình phương cực tiểu**

Giả sử b là 1 ước lượng bất kỳ của β. Khi đó tổng bình phương các sai lệch giữa các quan sát y_j và $b_0 + b_1x_{j1} + b_2x_{j2} + \dots + b_kx_{jk}$ – gọi là *giá trị ước lượng* của y_j với $j = \overline{1, n}$ là

$$S(b) = \sum_{j=1}^{n} \left(y_j - b_0 - b_1 x_{j1} - \dots - b_k x_{jk}\right)^2 = (y - Xb)^T(y - Xb) \quad (1.12)$$

Nội dung của phương pháp bình phương cực tiểu là tìm véc tơ $b = (b_0, b_1, \dots, b_n) \in R^{n+1}$ sao cho phiếm hàm S(b) đạt cực tiểu.

Véc tơ $\hat{\beta}$ làm cực tiểu phiếm hàm S(b) được gọi là *ước lượng bình phương cực tiểu* của β. Còn các giá trị

$$\hat{\varepsilon}_j = y_j - (\hat{\beta}_0 + \hat{\beta}_1 x_{j1} + \dots + \hat{\beta}_k x_{jk}); \; j = \overline{1, n}$$

gọi là các phần dư của phép hồi quy.

Đặt $\hat{y}_j = \hat{\beta}_0 + \hat{\beta}_1 x_{j1} + ... + \hat{\beta}_k x_{jk}$; $j = \overline{1,n}$; $\hat{y} = (\hat{y}_1; \hat{y}_2; ....; \hat{y}_n)^T$

Khi đó ta có

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + ... + \hat{\beta}_k X_k$$

được gọi là *phương trình hồi quy tuyến tính mẫu.*

1.5.2. Sự tồn tại của ước lượng bình phương cực tiểu (tính chất đại số của ước lượng bình phương cực tiểu)

Định lý 1.1 [1, 21, 35] Với các giả thiết (1.8) – (1.10) và $r(X) = k + 1 \le n$, thì ước lượng bình phương cực tiểu của bài toán hồi quy bội tuyến tính là tồn tại và là nghiệm của hệ Gauss – Markov:

$$(X^T X)\,\hat{\beta} = X^T y \qquad (1.13)$$

Hay $\hat{\beta} = (X^T X)^{-1} X^T y$

Chứng minh kết quả này có thể tham khảo ở [1, 21, 35]

Chú ý 1.1 Theo kết quả của định lý 1.1, ta có tổng bình phương các phần dư của mô hình (1.5):

$$\sum_{j=1}^{n} \hat{\varepsilon}_j^2 = S(\hat{\beta}) = (y - X\hat{\beta})^T (y - X\hat{\beta}) = y^T y - (y^T X)\hat{\beta}$$

Ký hiệu $\hat{\varepsilon} = (\hat{\varepsilon}_1; \hat{\varepsilon}_2; ....; \hat{\varepsilon}_n)^T$ là véc tơ phần dư của ước lượng bình phương cực tiểu của mô hình (1.7). Ký hiệu $M_X = I - X(X^T X)^{-1} X^T$ là ma trận vuông cấp n.

Ma trận M_X có tính chất

i) M_X đối xứng, tức là $M_X^T = M_X$.

ii) M_X luỹ đẳng, tức là $(M_X)^2 = M_X$.

iii) $X^T M_X = X^T(I - X(X^T X)^{-1} X^T) = X^T - X^T = 0$

Khi đó, ta có

$$\hat{\varepsilon} = y - X\,\hat{\beta} = [\, I - X\,(X^T X)^{-1} X^T\,]y = M_X y$$

Nên từ $M_X X = 0 \Rightarrow \hat{\varepsilon}^T . X = 0$

Mặt khác, ta lại có

$$\hat{\varepsilon} = M_X y = M_X\,(X\beta + \varepsilon) = M_X \varepsilon$$

Mối liên hệ giữa β và $\hat{\beta}$:

$$\hat{\beta} = (X^TX)^{-1}X^T(X\beta + \varepsilon) = \beta + (X^TX)^{-1}X^T\varepsilon$$

Chú ý 1.2. Các tính chất thống kê của ước lượng bình phương cực tiểu $\hat{\beta}$ đều phụ thuộc vào các giả thiết về véc tơ x_t và sai số ε (có thể tham khảo ở [21], trang 200 – 232).

1.6. Kỳ vọng có điều kiện

Trong toàn bộ mục này, ta luôn xét trong không gian xác suất cơ sở $(\Omega, \mathcal{A}, P)$.

1.6.1. Định nghĩa

Trước hết, ta giới thiệu khái niệm σ - trường.

Họ F các biến cố của Ω được gọi là σ - trường nếu

(i) $\Omega \in F$;

(ii) $A \in F \Rightarrow \Omega \backslash A \in F$;

(iii) $A_n \in F$; $n = 1, 2, 3, \ldots$ thì $\bigcup_{n=1}^{\infty} A_n \in F$.

1.6.1.1. Đối với phân hoạch

Giả sử $\{B_n\}$ là 1 phân hoạch của Ω, tức là $B_i \cap B_j = \phi; \forall i \neq j; \Omega = \sum_n B_n$

Ký hiệu F là σ - trường sinh bởi phân hoạch này (tức là lớp các tập có dạng $\bigcup_{n \in I} B_n$, I là tập con của $\{1, 2, 3, \ldots..\}$). Giả sử $P(B_n) > 0, \forall n$.

+) Với $A \in \mathcal{A}$, đặt $P(A/F) = \sum_n P(A/B_n)I_{B_n}$

trong đó $I_{B_n}(x) = 1, \forall x \in B_n; I_{B_n}(x) = 0, \forall x \notin B_n$;

+) Với X là biến ngẫu nhiên có kỳ vọng, tức là min $\{EX^+, EX^-\} < \infty$, đặt

$$E(X/F) = \sum_n E(X/B_n)I_{B_n}$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)};\ E(X/B) = \frac{1}{P(B)}\int_B XdP;\ \forall B \in F.$$

Như vậy, $E(X/F) = E(X/B_n)$ trên B_n. Từ đó suy ra $E(X/F)$ là biến ngẫu nhiên F - đo được và $\int_B E(X/F)dP = \int_B XdP; \forall B \in F$.

1.6.1.2. Đối với σ - trường

Giả sử F là σ - trường con của $\mathcal{A}$.

* Kỳ vọng có điều kiện của biến ngẫu nhiên $X \geq 0$ đối với F là biến ngẫu nhiên không âm $E(X/F)$: $\Omega \to [0, \infty)$ sao cho

(i) $E(X/F)$ là F - đo được.

(ii) Với mọi $A \in F$: $\int_A XdP = \int_A E(X/F)dP$

* X là biến ngẫu nhiên bất kỳ sao cho $\min\{E(X^+/F), E(X^-/F)\} < \infty$ với xác suất 1. Khi đó, nói X có kỳ vọng có điều kiện đối với σ - trường F và gọi $E(X/F) = E(X^+/F) - E(X^-/F)$ là kỳ vọng có điều kiện của X đối với F.

Đặc biệt, nếu $E(|X|) < \infty$ thì kỳ vọng có điều kiện của X đối với F là biến ngẫu nhiên có kỳ vọng hữu hạn là $E(X/F)$: $\Omega \to [0, \infty)$ được xác định bởi 2 điều kiện (i) và (ii).

* Phương sai có điều kiện:

$$Var(X/F) = E[(X - E(X/F))^2 / F]$$

* Kỳ vọng có điều kiện của X đối với biến ngẫu nhiên Y:

$$E(X/Y) = E(X/\sigma(Y))$$

Với $\sigma(Y)$ là σ - trường $Y^{-1}(B(R))$; $B(R)$ là σ - trường Borel trên R.

* Kỳ vọng có điều kiện đối với véc tơ ngẫu nhiên $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$:

$$E(X/\mathbf{Y}) = E(X/\sigma(\mathbf{Y}))$$

Với $\sigma(\mathbf{Y})$ là σ - trường $\mathbf{Y}^{-1}(B(R^n))$; $B(R^n)$ là σ - trường Borel trên R^n.

1.6.2. Tính chất của kỳ vọng có điều kiện

Từ định nghĩa của kỳ vọng có điều kiện, ta có thể chứng minh được các tính chất sau.

Định lý 1.2 [4, 35] Giả thiết rằng với các biến ngẫu nhiên bất kỳ, kỳ vọng có điều kiện luôn tồn tại và các quan hệ ở đây được hiểu là hầu chắc chắn, thì

(1) Nếu X là F - đo được thì $E(X/F) = X$.

Đặc biệt, nếu c là hằng số thì $E(c/F) = c$

(2) $E(aX + bY/F) = a\,E(X/F) + bE(Y/F)$ với $\forall a, b \in R$.

(3) Nếu $X \le Y$ thì $E(X/F) \le E(Y/F)$. Đặc biệt, ta có

$$|E(X/F)| \le E(|X|/F)$$

(4) $E(E(X/F)) = EX$

(5) Nếu $\sigma(X)$ và F là độc lập thì $E(X/F) = EX$.

Đặc biệt, nếu X, Y độc lập thì $E(X/Y) = EX$.

(6) Nếu $F_1 \subseteq F_2$ thì

$E[\,E(X/F_2)/F_1] = E[\,E(X/F_1)/F_2] = E(X/F_1)$.

(7) Nếu Y là F - đo được thì $E(XY/F) = Y.\,E(X/F)$

(8) Nếu $F_0 = \{\phi, \Omega\}$ thì $E(X/F_o) = EX$

(9) Nếu $g: R \to R$ là hàm lồi, tức là $g(ax + by) \le ag(x) + bg(y)$; $0 \le a, b \le 1$, $a + b = 1$; $x, y \in R$ thì $g(E(X/F) \le E(g(X)/F)$.

Chứng minh các kết quả này có thể tham khảo ở [4, 35] □

1.6.3. Không gian $L^2(\Omega, A, P)$

Xét không gian xác suất $(\Omega, \mathcal{A}, P)$ và lớp

$$\mathcal{C} = \left\{X \in \Omega : EX^2 := \int_\Omega X^2(\omega)P(d\omega) < \infty\right\}$$

Khi đó, ta có $E(aX)^2 = a^2\,EX^2$; $\forall a \in R, \forall X \in \mathcal{C}$ (1.16)

Mặt khác $(X + Y)^2 \le 2\,X^2 + 2\,Y^2$ nên

$$E(X + Y)^2 \le 2\,EX^2 + 2\,EY^2;\ \forall X, Y \in \mathcal{C} \quad (1.17)$$

Từ (1.16) và (1.17) ta có $aX \in \mathcal{C}$; $\forall X \in \mathcal{C}$; $\forall a \in R$ và $X + Y \in \mathcal{C}$; $\forall X, Y \in \mathcal{C}$.

Do vậy, $\mathcal{C}$ là 1 không gian véc tơ trên R với 2 phép toán cộng 2 biến ngẫu nhiên và nhân số thực với biến ngẫu nhiên.

1.6.3.1. Tích vô hướng trong $\mathcal{C}$

$\forall$ X, Y $\in \mathscr{C}$; ta định nghĩa

$$<X, Y> = E(XY)$$

Khi đó, dễ dàng kiểm tra <X, Y> thoả mãn đầy đủ các tính chất của tích vô hướng thông thường trong các không gian Euclide [3]. Riêng <X, X > = 0 không suy ra $X(\omega) = 0\ \forall\ \omega \in \Omega$ mà chỉ có $P(X(\omega) = 0) = 1$. Để khắc phục điều này, định nghĩa X tương đương với Y nếu $P(X(\omega) = Y(\omega)) = 1$; ký hiệu X ~ Y. Đây là 1 quan hệ tương đương trên $\mathscr{C}$ và lớp tương đương ký hiệu [X] = X/ ~ . Khi đó, định nghĩa

$$L^2(\Omega, \mathcal{A}, P) = \{ [X]: X \in \mathscr{C} \}$$

Trang bị cho $L^2(\Omega, \mathcal{A}, P)$ chuẩn: $\|X\| = \sqrt{<X,X>}$. Khi đó ($L^2(\Omega, \mathcal{A}, P)$, $\|\ \|$) là 1 không gian tuyến tính định chuẩn.

Để đơn giản, từ nay trở về sau ta dùng X thay cho [X], L^2 thay cho $L^2(\Omega, \mathcal{A}, P)$.

1.6.3.2. Sự hội tụ trong không gian L^2

Định nghĩa 1.1 Dãy $\{X_n\}$, $X_n \in L^2$ được gọi là hội tụ tới X nếu $\|X_n - X\|^2 = E|X_n - X|^2 \to 0$ khi $n \to \infty$; ký hiệu $X_n \xrightarrow{ms} X$.

Sự hội tụ này được gọi là *hội tụ theo trung bình bình phương.*

Từ định nghĩa ta dễ dàng chứng minh được các tính chất sau:

Định lý 1.4 [3] Với không gian $L^2 = L^2(\Omega, \mathcal{A}, P)$ cùng với sự hội tụ theo trung bình bình phương trên đó, thì ta có:

(1) Không gian L^2 là không gian đầy hay L^2 là không gian Hilbert.

(2) $X_n \xrightarrow{ms} X \Leftrightarrow E|X_n - X_m|^2 \to 0$ khi $m, n \to \infty$.

Chứng minh các kết quả này có thể tham khảo ở [3] □

1.6. 3.3. Định lý chiếu trong không gian Hilbert

Định lý 1.5 [7] Cho M là 1 không gian con đóng của không gian Hilbert H và x $\in$ H thì

(i) tồn tại duy nhất $\hat{x} \in M$ sao cho

$$\|x-\hat{x}\| = \inf_{y\in M}\|x-y\|$$

(ii) $\hat{x} \in M$ và $\|x-\hat{x}\| = \inf_{y\in M}\|x-y\|$ khi và chỉ khi $\hat{x} \in M$

và $(x-\hat{x}) \in M^{\perp}$

Trong đó $M^{\perp}$ là không gian con trực giao với M

Khi đó $\hat{x}$ được gọi là chiếu trực giao của x lên M và ký hiệu $\hat{x} = P_M x$.

Chứng minh các kết quả này có thể tham khảo ở [7] □

Định nghĩa 1.2 Toán tử $P_M : H \to M$ xác định bởi $\hat{x} = P_M x$, $\forall x \in H$ gọi là phép chiếu trực giao từ H lên M.

Từ định lý và định nghĩa phép chiếu trực giao, dễ dàng kiểm tra được các tính chất sau [7]:

(1) $P_M(\alpha x + \beta y) = \alpha P_M x + \beta P_M y$; $\forall x, y \in H$ và $\alpha, \beta \in R$. Điều này suy ra toán tử P_M là 1 toán tử tuyến tính.

(2) $\forall x \in H$ thì $\|x\|^2 = \|P_M x\|^2 + \|(I-P_M)x\|^2$ với I là toán tử đồng nhất trên H. Tính chất này suy ra P_M là toán tử tuyến tính liên tục.

(3) $\forall x \in H$ thì x viết được dưới dạng duy nhất:

$$x = P_M x + (I-P_M)x \text{ với } P_M x \in M \text{ và } (I-P_M)x \in M^{\perp}.$$

(4) $P_M x_n \to P_M x \Leftrightarrow \|x_n - x\| \to 0$ khi $n \to \infty$

(5) $x \in M \Leftrightarrow P_M x = x$

(6) $x \in M^{\perp} \Leftrightarrow P_M x = 0$

(7) Nếu $M_1 \subseteq M_2$ thì $P_{M_1}(P_{M_2}x) = P_{M_1}x$; $\forall x \in H$. Đặc biệt $P_M^2 = P_M$

Chú ý 1.3 đặt $Q = I - P_M$ thì Q chính là phép chiếu trực giao của H lên $M^{\perp}$.

1.6.3.4. Kỳ vọng có điều kiện như phép chiếu trong không gian Hilbert L^2

Giả sử F là σ - trường con của $\mathcal{A}$, gọi L_F^2 là không gian con các phần tử F - đo được của L^2. Khi đó $L_F^2 = \{ X \in L^2 : \sigma(X) \subseteq F\}$. Xác định toán tử

$$P_F X = E(X/F) \text{ với } \forall X \in L^2$$

Định lý 1.6 [35] Toán tử P_F có các tính chất sau

(1) P_F là toán tử tuyến tính trên L^2;

(2) $P_F^2 = P_F$. Vì vậy, P_F chính là phép chiếu trực giao từ L^2 vào L_F^2;

(3) $I - P_F$ là trực giao với L_F^2;

(4) $\forall Y \in L^2, Z \in L_F$: $<P_FY, Z> = <Y, P_FZ>$;

(5) $\|P_FX\| \le \|X\|$ và $\|P_F\| = 1$ nếu $\dim(L_F) \ge 1$.

Ở đây, $\|P_F\| = \sup\limits_{\{Y \in L^2 : Y \neq 0\}} \dfrac{\|P_FY\|}{\|Y\|}$.

Chứng minh các kết quả này có thể tham khảo ở [35] □

Chú ý 1.4 Từ định lý trên, ta có E(X/ F) chính là phép chiếu trực giao của L^2 vào L_F^2 nên theo định lý hình chiếu E(X/F) chính là ước lượng tốt nhất của X theo nghĩa

$$E[X - E(X/F)]^2 = \inf\{E[X - Y]^2 : Y \in L_F^2\}$$

1.6.3. 5. Phương trình dự đoán

Cho không gian Hilbert L^2, tập con đóng $M \subseteq L^2$ và $X \in L^2$, định lý hình chiếu trong L^2 khẳng định tồn tại duy nhất $\hat{X} \in M$ sao cho

$\|X - \hat{X}\|^2$ là nhỏ nhất

$$\Leftrightarrow <X - \hat{X}, Y> = 0 \ \forall Y \in M \tag{1.18}$$

Phương trình (1.18) được gọi là *phương trình dự đoán* và phần tử $\hat{X} = P_MX$ gọi là dự đoán tốt nhất của X trong M. Nói cách khác, dự đoán tốt nhất của X trong M chính là chiếu trực giao của X lên M. Từ định nghĩa phép chiếu trong L^2, ta có

$$\|X - \hat{X}\|^2 = \inf_{Z \in M}\|X - Z\|^2 = \inf_{Z \in M} E|X - Z|^2 \tag{1.19}$$

Điều này có nghĩa rằng $\hat{X}$ là *dự báo trung bình bình phương tốt nhất* của X trong M và đó chính là $P_M X$ và theo (1.19) cũng chính là E(X/M).

Dự báo tuyến tính tốt nhất

Nếu $M = \text{span}\{X_k, X_k \in L^2, k \in I; I \subset N\}$ thì $\hat{X} = P_MX$ được gọi là *dự đoán tuyến tính tốt nhất* của X theo các thành phần của $\{X_k, X_k \in L^2, k \in I\}$.

CHƯƠNG 2. CÁC MÔ HÌNH CHUỖI THỜI GIAN TÀI CHÍNH DẠNG ARCH

Trong lý thuyết chuỗi thời gian của Box – Jenskin [15], giả thiết dừng là giả thiết bản chất cho các mô hình ARMA, các phương pháp ước lượng tham số của mô hình phụ thuộc vào giả thiết này và giả thiết về sai số của mô hình $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ hoặc $IID(0, \sigma^2)$, hoặc $N(0, \sigma^2)$ (các khái niệm này sẽ giới thiệu ở mục 2.2). Nội dung chính của chương này của đề tài tập trung vào vấn đề nhận dạng 2 mô hình: ARMA và từ nhận dạng mô hình ARMA để nhận dạng mô hình ARCH nhằm giải quyết 2 bài toán cơ bản của phân tích chuỗi thời gian:

- Nhận dạng mô hình chuỗi thời gian.
- Kiểm định sự phù hợp của mô hình chuỗi thời gian.

Đối với mô hình ARCH/ GARCH, mô hình mà phương sai có điều kiện thay đổi theo thời gian thay vì bằng hằng số như mô hình ARMA, điều này giải thích cho hiện tượng các chuỗi thời gian tài chính có phân bố nặng đuôi. Để có thể giải thích được những hiện tượng này, ta tiếp cận phân tích theo 3 hướng:

- Kiểm định nhằm phát hiện hiệu ứng ARCH/ GARCH trên chuỗi thời gian.
- Ước lượng tham số của mô hình ARCH/ GARCH.
- Kiểm định nhằm phát hiện sự biến mất của hiệu ứng ARCH/ GARCH

Trong đó, chú trọng bài toán ước lượng tham số của mô hình ARCH/ GARCH bằng phương pháp bình phương cực tiểu.

2.1. Quá trình dừng

2.1.1. Chuỗi thời gian

Định nghĩa 2.1 *Chuỗi thời gian* là một dãy các giá trị quan sát $\mathcal{X} = \{x_t;\ t = \overline{1, n}\}$ được xếp theo thứ tự diễn biến thời gian. Trong đó x_t là quan sát tại thời điểm t (giây, phút, giờ, ngày, tháng, năm ,...) hay $t \in T$; T là tập rời rạc. Trong đề tài chúng ta coi dãy quan sát $\mathcal{X}$ là một chuỗi thời gian và đồng thời cũng là thể hiện của một quá trình ngẫu nhiên $\{X_t, t \in T\}$ xác định trên không gian xác suất $(\Omega, \mathcal{A}, P)$ [3, 4, 5]. Từ đây trở về sau khi nói tới quá trình ngẫu nhiên $\{X_t\}$, ta hiểu $t \in Z$.

Mục đích nghiên cứu của chuỗi thời gian là từ thông tin của tập quan sát, tìm kiếm các kết luận về các tính chất và nét bản chất của quá trình ngẫu nhiên. Từ đó tìm cách xây dựng mô hình phù hợp với chuỗi quan sát đó và sử dụng mô hình đó để dự báo. Một trong những công cụ để nghiên cứu chuỗi thời gian là quá trình dừng.

2.1.2. Quá trình dừng

Trong đề tài, quá trình dừng được hiểu theo nghĩa sau:

Định nghĩa 2.2 Giả sử $\{X_t\}$ là quá trình có $Var(X_t) < \infty$. $\{X_t\}$ được gọi là *quá trình dừng* nếu hàm $\mu_X(t) = E(X_t)$ $(t \in Z)$ là hằng số (không phụ thuộc vào t) và hàm tự hiệp phương sai

$$\gamma_X(r,s) = cov(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))] \quad (r,s \in Z)$$

chỉ phụ thuộc vào $s - r$.

Như vậy, $\{X_t\}$ là quá trình dừng khi và chỉ khi

a) $E|X_t|^2 < \infty,\ t \in Z$;

b) $\mu_X(t) = m,\ \forall\ t \in Z$;

c) $\gamma_X(r,s) = \gamma_X(r+t, s+t);\ \forall\ t, r, s \in Z$.

Nói cách khác, quá trình $\{X_t\}$ là *quá trình dừng* nếu nó cùng hàm trung bình và hàm tự tương quan với quá trình $\{X_{t+h}\}$, $\forall h \in Z$.

*** Hàm tự hiệp phương sai của quá trình dừng**

Cho $\{X_t\}$ là quá trình dừng.

Định nghĩa 2.3 Hàm tự hiệp phương sai của quá trình dừng $\{X_t\}$ (Autocovariace function – ACVF) xác định bởi $\gamma_X(h) = Cov(X_{t+h}, X_t)$.

Hàm tự tương quan của $\{X_t\}$ (Autocorrelation function – ACF) xác định bởi

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = Corr(X_{t+h}, X_t)$$

Định lý 2.1 [4] (Các tính chất của hàm tự hiệp phương sai của quá trình dừng)

i) $\gamma_X(h)$ là hàm chẵn.

ii) $\gamma_X(0) \geq 0$ và $|\gamma_X(h)| \leq \gamma_X(0), \forall h \in Z$;

iii) $\gamma_X(h)$ là xác định không âm, tức là $\forall n \in N$; $\forall t_1, t_2, \ldots, t_n \in Z$; $\forall b_1, b_2, \ldots, b_n \in R$ thì $\sum_{i=1}^{n}\sum_{j=1}^{n} b_i b_j \gamma_X(t_i - t_j) \geq 0$.

Chứng minh định lý 2.1 có thể tham khảo ở [4] □

Để đơn giản trong trình bày, đối với quá trình dừng $\{X_t\}$ sử dụng ký hiệu $\gamma_h; \rho_h$ lần lượt thay cho $\gamma_X(h); \rho_X(h)$.

Trong thực tế, ta chỉ quan sát được một thể hiện hữu hạn $\mathscr{X} = \{x_t;\ t = \overline{1, n}\}$ của một chuỗi thời gian dừng, do đó về nguyên tắc, ta không biết được chính xác các hàm tự hiệp phương sai của chuỗi thời gian đó. Vì thế, từ thể hiện đó ta muốn ước lượng hàm tự hiệp phương sai của chuỗi thời gian $\{X_t\}$, đó là bước quan trọng trong việc xây dựng mô hình toán học cho dữ liệu. Ước lượng của hàm $\gamma_X(h)$ chính là hàm tự hiệp phương sai mẫu.

Cho $\mathscr{X} = \{x_t;\ t = \overline{1, n}\}$ là 1 thể hiện của chuỗi thời gian $\{X_t\}$.

Định nghĩa 2.4

Trung bình mẫu của $\mathscr{X}$ là:

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$

Hàm tự hiệp phương sai mẫu của $\mathscr{X}$ là:

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-|h|}(x_{t+|h|} - \bar{x})(x_t - \bar{x}) \text{ với } -n < h < n$$

Hàm tự tương quan mẫu là:

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \text{ với } -n < h < n$$

Chú ý 2.1 Các tính chất về dáng điệu tiệm cận của $\bar{x}, \hat{\gamma}, \hat{\rho}$ có thể tham khảo ở [17].

2.2. Quá trình tự hồi quy trung bình trượt ARMA

(AutoRegressive Moving Average)

2.2.1. Ồn trắng

Định nghĩa 2.5 Quá trình dừng $\{\varepsilon_t\}$ với kỳ vọng 0 và phương sai σ^2 được gọi là *ồn trắng*, ký hiệu $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ nếu và chỉ nếu:

i) $\{\varepsilon_t\}$ có kỳ vọng 0;

ii) Hàm tự hiệp phương sai:

$$\gamma_\varepsilon(h) = \begin{cases} \sigma^2 & \text{khi } h = 0 \\ 0 & \text{khi } h \neq 0 \end{cases}$$

* Trong trường hợp, $\{\varepsilon_t\}$ độc lập và có cùng phân bố với với kỳ vọng 0 và phương sai σ^2, thì ký hiệu $\{\varepsilon_t\} \sim IID(0, \sigma^2)$.

2.2.2. Quá trình tự hồi quy AR(p) (AutoRegressive)

Định nghĩa 2.6 Quá trình $\{X_t\}$ là quá trình nhân quả tự hồi quy cấp p, ký hiệu $\{X_t\} \sim AR(p)$ nếu $\{X_t\}$ là quá trình dừng thoả mãn:

$$(\forall t \in Z) : X_t - a_1X_{t-1} - a_2X_{t-2} - \ldots - a_pX_{t-p} = \varepsilon_t \,;\; a_p \neq 0 \quad (2.1)$$

Hay $a(B)\, X_t = \varepsilon_t$

Trong đó B là toán tử xác định trên không gian $(\Omega, \mathcal{A}, P)$:

$$B_iX_t = X_{t-i};\; i = 1,2,3, \ldots \quad (2.2)$$

$\{\varepsilon_t\} \sim WN(0, \sigma^2)$, đa thức tự hồi quy a(z) là hàm xác định bởi

$a(z) = 1 - a_1z - a_2z^2 - \ldots - a_p z^p$

với a(z) có nghiệm nằm ngoài đĩa tròn đơn vị trong mặt phẳng phức.

2.2.3. Quá trình trung bình trượt MA(q) (Moving Average)

Định nghĩa 2.7 Quá trình $\{X_t\}$ là quá trình trung bình trượt cấp q, ký hiệu $X_t \sim MA(q)$ nếu $\{X_t\}$ là quá trình thoả mãn $(\forall t \in Z)$:

$$X_t = \varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t2} + \ldots + b_q\varepsilon_{t-q};\; b_q \neq 0 \quad (2.3)$$

Hay $X_t = b(B)\, \varepsilon_t$

Trong đó $\{\varepsilon_t\} \sim WN(0, \sigma^2)$, đa thức tự hồi quy b(z) là hàm xác định bởi

$b(z) = 1 + b_1z + b_2z^2 + \ldots + b_q z^q$

2.2.4. Quá trình tự hồi quy trung bình trượt ARMA(p, q)

Định nghĩa 2.8 Quá trình $\{X_t\}$ là quá trình tự hồi quy trung bình trượt cấp (p, q), ký hiệu $\{X_t\}$ ~ ARMA(p, q) nếu $\{X_t\}$ là 1 quá trình dừng thoả mãn:

$$(\forall t \in Z) : X_t - a_1X_{t-1} - a_2X_{t-2} - \dots - a_pX_{t-p}$$
$$= \varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t2} + \dots + b_q\varepsilon_{t-q}; a_p, b_q \neq 0 \qquad (2.4)$$

Hay $a(B)X_t = b(B)\varepsilon_t$, $\{\varepsilon_t\} \sim WN(0, \sigma^2)$.

Đa thức tự hồi quy a(z) là hàm xác định bởi:

$$a(z) = 1 - a_1z - a_2z^2 - \dots - a_p z^p$$

đa thức trung bình trượt b(z) là hàm xác định bởi:

$$b(z) = 1 + b_1z + b_2z^2 + \dots + b_q z^q$$

Các đa thức $a(z) = 1 - \sum_{i=1}^{p} a_i z^i$, $b(z) = \sum_{j=0}^{q} b_j z^j$; $b_o = 1$ có tất cả các nghiệm nằm ngoài đĩa tròn đơn vị trong mặt phẳng phức.

2.3. Qúa trình tự hồi quy với những biến động bất thường – ARCH(Q) (AutoRegresive Conditional Heteroskedasticity)

Các mô hình chuỗi thời gian ARMA(p, q) trình bày ở 2.2 chỉ thành công trong dự báo kỳ vọng và thất bại trong dự báo phương sai, đó là lớp các chuỗi thời gian tài chính. ARCH viết tắt cho *AutoRegresive Conditional Heteroskedasticity.* Thuật ngữ Heteroskedasticity được giải thích là hiện tượng bất thường về phương sai mà nguyên nhân chủ yếu là do các quá trình ngẫu nhiên bên ngoài tác động vào.

Mô hình ARCH là cống hiến mang tính chất khai sáng của Engle [19]. Qua nhiều năm nghiên cứu chuỗi thời gian tài chính ông thấy phương sai của chuỗi thời gian biểu hiện dưới 2 hình thức: dạng có điều kiện (ngắn hạn) và dạng không điều kiện (dài hạn). Lần đầu tiên ông xây dựng thành công mô hình giải thích những bất thường của phương sai mà chỉ sử dụng thông tin quá khứ của bản thân nhiễu. Conditional trong tên gọi ARCH có nghĩa là dựa vào thông tin quá khứ của bản thân quá trình. Một kết quả đáng lưu ý của mô hình ARCH là sự tiệm cận của phương sai có điều kiện về phương sai không điều kiện.

2.3.1. Định nghĩa

Giả sử $\{\varepsilon_t\}$ là quá trình ngẫu nhiên nhận giá trị rời rạc và F_t là tập thông tin quá khứ (σ- trường) tạo bởi các giá trị ε_t, ε_{t-1}, ... Khi đó quá trình ARCH(Q) (AutoRegresive Conditional Heteroskedasticity) được định nghĩa bởi

$$\varepsilon_t\big|_{F_{t-1}} \sim IID(0, h_t) \Leftrightarrow \varepsilon_t = \upsilon_t.\sqrt{h_t}\text{; với } \{\upsilon_t\} \sim IID(0,1) \qquad (2.5)$$

$$h_t = \alpha_o + \sum_{i=1}^{Q} \alpha_i \varepsilon_{t-i}^2 \qquad (2.6)$$

Ở đây Q > 0, Q nguyên dương

$$\alpha_0 > 0;\ \alpha_i \geq 0, i = \overline{1,Q} \qquad (2.7)$$

$$\sum_{i=1}^{Q} \alpha_i < 1 \qquad (2.8)$$

Trong thực hành, thường giả thiết $\{\upsilon_t\} \sim N(0,1)$.

Ký hiệu $\{\varepsilon_t\} \sim ARCH(Q)$.

Thông thường $\{\varepsilon_t\}$ là sai số của mô hình hồi quy tuyến tính:

$$y_t = x_t^T.\beta + \varepsilon_t;\ t = \overline{1,n}$$

Trong x_t là véc tơ k – chiều chỉ biến độc lập, $Y = (y_1, y_2, ..., y_n)^T$ véc tơ chỉ biến giải thích, β là véc tơ tham số của mô hình, x_t có thể chứa các giá trị trễ của biến giải thích. Từ đây, ta thấy mối liên hệ giữa nhiễu $\{\varepsilon_t\}$ và phương sai có điều kiện theo mô hình (2.5) và (2.6) là mối liên hệ mang tính bản chất và là giả thiết chung cho các mô hình kiểu ARCH. Mô hình (2.6) giải thích phương sai có điều kiện thay đổi theo thời gian và phụ thuộc vào chính nhiễu $\{\varepsilon_t\}$ theo mô hình hồi quy.

Trong trường hợp đặc biệt, $\alpha_i = 0\ \forall i = \overline{1,Q} \Rightarrow h_t = \alpha_o$ nên $\{\varepsilon_t\}$ là 1 quá trình có $Var(\varepsilon_t / F_{t-1}) = \alpha_o$

Giả thiết (2.7) đảm bảo cho phương sai có điều kiện là số dương. Còn giả thiết (2.8) đảm bảo cho quá trình $\{\varepsilon_t\}$ có mômen cấp 2 hữu hạn.

2.3.2. Tính chất của quá trình ARCH(Q)

Giả sử $\{\varepsilon_t\}$ ~ ARCH(Q), sau đây là các đặc trưng cơ bản của quá trình ARCH(Q).

2.3.2.1. Các đặc trưng không điều kiện

* Kỳ vọng không điều kiện:

$$E(\varepsilon_t) = E\left(\upsilon_t\sqrt{\alpha_0 + \sum_{i=1}^{Q}\alpha_i\varepsilon_{t-i}^2}\right) = E(\upsilon_t).E\left(\sqrt{\alpha_0 + \sum_{i=1}^{Q}\alpha_i\varepsilon_{t-i}^2}\right) = 0$$

Vậy $E(\varepsilon_t) = 0$

* $E(\varepsilon_t\,\varepsilon_{t-s}) = 0$ nếu $s \neq 0$. Điều này suy ra từ $E(\upsilon_t\upsilon_{t-s}) = 0$. Nên $\{\varepsilon_t\}$ không tương quan.

* Phương sai không điều kiện:

$$Var(\varepsilon_t) = E(\varepsilon_t^2) = E(h_t\upsilon_t^2) = E[E(h_t\upsilon_t^2 / F_{t-1})] = E[h_tE(\upsilon_t^2 / F_{t-1})]$$

$$= E(h_t) = E(\alpha_o + \sum_{i=1}^{Q}\alpha_i\varepsilon_{t-i}^2) \Rightarrow \; Var(\varepsilon_t) = \frac{\alpha_o}{1 - \sum_{i=1}^{Q}\alpha_i}$$

Từ công thức này, ta thấy điều kiện (2.8) ta áp cho các tham số của mô hình là hợp lý vì nó đảm bảo cho phương sai không điều kiện dương.

2.3.2.2. Tính chất phân bố nặng đuôi của quá trình ARCH

Để đơn giản xét quá trình ARCH(1). Ta tính moment cấp 4 của quá trình đó. $m_4 = E(\varepsilon_t^2)^2 = \dfrac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$. Kết quả này suy ra

+) Môment cấp 4 là dương $\Leftrightarrow$ $0 \le \alpha_1^2 < \dfrac{1}{3}$

+) Độ nhọn của $\{\varepsilon_t\}$: $\dfrac{m_4}{[Var(\varepsilon_t)]^2} = 3.\dfrac{1-\alpha_1^2}{1-3\alpha_1^2} > 3$

Điều này giải thích quá trình ARCH(1) có phân bố nặng đuôi hơn phân bố chuẩn. *Kết quả này giải thích cho hiện tượng nặng đuôi trong các chuỗi thời gian tài* *chính. Điều này minh chứng cho sự xuất hiện hiệu ứng ARCH trên các chuỗi thời gian tài chính.*

Các kết quả này vẫn đúng cho quá trình ARCH(Q) tổng quát, tuy nhiên việc tính toán cũng hơi phức tạp, trong khuôn khổ đề tài này không đề cập tới.

2.3.2.2. Các đặc trưng có điều kiện

* Kỳ vọng:

$$E(\varepsilon_t / F_{t-1}) = E(\upsilon_t . \sqrt{\alpha_o + \sum_{i=1}^{Q} \alpha_i \varepsilon_{t-i}^2} / F_{t-1}) = \sqrt{\alpha_o + \sum \alpha_i \varepsilon_{t-i}^2} . E(\upsilon_t / F_{t-1}) = 0$$

vì υ_t độc lập với các giá trị quá khứ $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$

* Phương sai:

$$Var(\varepsilon_t / F_{t-1}) = E(\varepsilon_t^2 / F_{t-1}) = E\left(\upsilon_t^2 .(\alpha_o + \sum_{i=1}^{Q} \alpha_i \varepsilon_{t-i}^2) / F_{t-1}\right)$$

$$= (\alpha_o + \sum_{i=1}^{Q} \alpha_i \varepsilon_{t-i}^2) E(\upsilon_t^2 . / F_{t-1}) = \alpha_o + \sum_{i=1}^{Q} \alpha_i \varepsilon_{t-i}^2 \qquad (2.9)$$

Công thức (2.9) giải thích hiện tượng phương sai có điều kiện khác hằng số trong khi phương sai không điều kiện là hằng số. Phương sai có điều kiện h_t phụ thuộc vào chính bình phương của quá trình theo mô hình hồi quy (2.9).

2.3.3. Kiểm định hiệu ứng ARCH

Xử lý thô chuỗi thời gian, tính toán các hàm tự tương quan và biểu diễn trên đồ thị, thực hiện các kiểm định nhằm phát hiện những biểu hiện của hiệu ứng ARCH trên chuỗi quan sát. Ở bước này, ta xây dựng một mô hình phù hợp với dữ liệu để loại bỏ bất kỳ sự phụ thuộc tuyến tính trong dữ liệu và dùng chuỗi phần dư của mô hình để kiểm định hiệu ứng ARCH. Đề tài sử dụng một trong hai phương pháp sau để kiểm tra hiệu ứng ARCH trên các chuỗi thời gian tài chính.

2.3.3.1. Kiểm định Ljung – Box

Mục đích của kiểm định Ljung – Box [24] là kiểm định sự phù hợp của mô hình ARMA, tức kiểm định $\{\varepsilon_t\} \sim IID(0, 1)$ không ?

Thủ tục kiểm định theo các bước sau:

Bước 1: Ước lượng chuỗi thời gian bằng mô hình ARMA tốt nhất và điều này được giới thiệu ở 2.3. Nhiễu thu được lấy bình phương. Tính toán các tự tương quan mẫu của nhiễu lấy bình phương.

$$\hat{\rho}_i = \frac{\sum_{t=i+1}^{n}(\hat{\varepsilon}_t^2 - \hat{\sigma}^2)(\hat{\varepsilon}_{t-i}^2 - \hat{\sigma}^2)}{\sum_{t=1}^{n}(\hat{\varepsilon}_t^2 - \hat{\sigma}^2)^2}; i = \overline{1,n} \text{ và } \hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_t$$

Bước 2: Tính thống kê

$$Q(h) = n(n+2)\sum_{i=1}^{h}\frac{\hat{\rho}_i^2}{n-i}$$

Kiểm định cặp giả thuyết:

H_o: không có hiệu ứng ARCH.

H_1: Có hiệu ứng ARCH.

Với giả thuyết H_o, n đủ lớn (n > 30) và h < n/2, Q tiệm cận tới phân phối χ_h^2. Với mức ý nghĩa α nếu Q > $\chi_{h,1-\alpha}^2$ thì bác bỏ giả thuyết gốc và chấp nhận đối thuyết.

2.3.3.2. Kiểm định nhân tử Lagrange

Kiểm định nhân tử Lagrange tiến hành ước lượng tham số cho mô hình ARCH(Q). Thay vì dùng thống kê nR^2 [36] để kiểm tra có chấp nhận giả thuyết khác không của các tham số, ta dùng thống kê F. Kiểm định này trải qua 3 bước:

Bước 1: Sử dụng phương pháp bình phương cực tiểu ước lượng tham số của mô hình AR(p) thích hợp:

$$X_t = a_o + a_1X_{t-1} + a_2X_{t-2} + \ldots + a_pX_{t-p} + \varepsilon_t$$

Bước 2: Nhiễu thu được bình phương và ước lượng mô hình AR(Q):

$$\varepsilon_t^2 = \alpha_o + \alpha_1\varepsilon_{t-1}^2 + \ldots + \alpha_Q\varepsilon_{t-Q}^2 + \eta_t;\ t = \overline{Q+1,n}$$

Bước 3: Kiểm định hiệu ứng ARCH:

Đặt $SSR_0 = \sum_{t=Q+1}^{n}\left(\hat{\varepsilon}_t^2 - \hat{\sigma}\right)^2$ và $SSR_1 = \sum_{t=Q+1}^{n}\hat{\eta}_t^2$. Trong đó $\hat{\varepsilon}_t;\hat{\eta}_t$ lần lượt là nhiễu ước lượng được từ mô hình AR(p), AR(Q) ở bước 1 và bước 2 và $\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_t$.

Sử dụng thống kê

$F = \dfrac{(SSR_0 - SSR_1)/Q}{SSR_1/(T-2Q-1)}$ trong kiểm định cặp giả thuyết

H_0: Không có hiệu ứng ARCH

$$\alpha_1 = \alpha_2 = ... = \alpha_Q = 0$$

H_1: Có hiệu ứng ARCH

$$\sum_{j=1}^{Q}\alpha_j^2 \neq 0$$

Với giả thuyết gốc, F tiệm cận tới phân phối χ_Q^2. Với mức ý nghĩa α, nếu F > $\chi_{Q,1-\alpha}$ thì bác bỏ giả thuyết gốc H_0 và chấp nhận đối thuyết H_1.

2.3.4. Ước lượng tham số cho mô hình ARCH(Q)

Các tham số của mô hình ARCH có thể ước lượng bởi nhiều phương pháp, bài toán ước lượng tham số của mô hình ARCH(Q) sử dụng trong đề tài dùng phương pháp bình phương cực tiểu sử dụng kỹ thuật Householder [3].

Viết lại mô hình ARCH(Q):

$$\varepsilon_t\big|_{F_{t-1}} \sim IID(0, h_t) \Leftrightarrow \varepsilon_t = \upsilon_t.\sqrt{h_t} \text{ với } \{\upsilon_t\} \sim IID(0,1)$$

$$h_t = \alpha_o + \sum_{i=1}^{Q}\alpha_i\varepsilon_{t-i}^2$$

Ta thêm các giả thiết $\{\upsilon_t\}$ có mô men cấp 4 hữu hạn và $\{\upsilon_t; t = \overline{1,n}\}$ là độc lập với $\{\varepsilon_t; 1-Q \le t \le 0\}$. Ngoài ra, để thu được các tính chất tiệm cận của các ước lượng ta giả sử quá trình $\{\varepsilon_t; 1-Q \le t\}$ là dừng và ergodic (giả thuyết về trung bình theo tập hợp và trung bình theo thời gian là trùng nhau).

Một phương pháp chung nhất dùng để ước lượng tham số của mô hình ARCH(Q) là tiếp cận theo ước lượng hợp lý Gauss. Theo cách này, ước lượng thu được là làm cực đại hàm log của hàm hợp lý Gauss, kết quả ước lượng gọi là *ước lượng tựa hợp lý cực đại Gauss* (QMLE). Ước lượng này là tồn tại thậm chí khi mật độ của của sai số có điều kiện không là phân bố chuẩn, sự tồn tại và tính chất tiệm cận chuẩn của ước lượng này được thiết lập bởi Weiss (1986, [23]). Các tính chất này vẫn có giá trị khi sai số là phân bố Gauss. Tuy nhiên các ước lượng hợp lý cực đại này rất khó có thể cài đặt tính toán bằng các chương trình máy tính, đặc biệt khi kích thước mẫu n tăng lên. Khắc phục nhược điểm này, đề tài trình bày bài toán ước lượng tham số của mô hình ARCH(Q) thông qua giải hai hệ thống phương trình tuyến tính mà không phải thực hiện thuật toán tối ưu phi tuyến. Kết quả chính (định lý 2.3) của ước lượng tham số của mô hình ARCH(Q) bằng phương pháp bình phương cực tiểu.

Để thu được ước lượng, đặt $y_t = \varepsilon_t^2, 1-Q \le t \le n; X_{t-1} = \left(1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-Q}^2\right)^T$;

$\alpha = (\alpha_0, \alpha_1, \dots, \alpha_Q)^T$ và $\eta_t = \upsilon_t^2 - 1$, thì

$$h_t(\alpha) = X_{t-1}^T \alpha \qquad (2.10)$$

và bình phương (2.5), dùng (2.10) ta có

$$y_t = X_{t-1}^T.\alpha + h_t(\alpha).\eta_t; \ t = \overline{1,n} \qquad (2.11)$$

Ở đây, $E(h_t(\alpha)\eta_t) = E(h_t(\alpha))E(\eta_t) = 0; \ t = \overline{1,n}$

Mô hình (2.11) có dạng mô hình hồi quy tuyến tính với sai số có trung bình không. Ở đây, ta bỏ qua tính ngẫu nhiên của $h_t(\alpha)$ và cũng như sự có mặt của α trong h_t. Khi đó ta thu được 1 ước lượng bình phương cực tiểu thô của α là

$$\hat{\alpha}_{pr} = (X^T X)^{-1} X^T Y \qquad (2.12)$$

Trong đó X là ma trận cấp n x (1 + Q) với dòng thứ i là véc tơ X_{i-1}^T

$(i = \overline{1,n})$ và $Y = \left(y_1; y_2; \dots y_n\right)^T$.

Ước lượng này không quan tâm đến phương sai có điều kiện khác hằng của các sai số và do đó nó chỉ là ước lương thô. Phân bố tiệm cận của ước lượng có thể thu được dễ dàng bằng định lý giới hạn trung tâm xấp xỉ và ước lượng này kém hiệu quả hơn ước lượng QMLE [23].

Sử dụng ước lượng (2.12) để cải tiến ước lượng $\hat{\alpha}_{pr}$ của α như sau:

Chia 2 vế của (2.11) cho $h_t(\alpha)$, thu được

$$\frac{y_t}{h_t(\alpha)} = \left(\frac{X_{t-1}}{h_t(\alpha)}\right)^T .\alpha + \eta_t$$

Trong biểu diễn này nếu chúng ta thay thế $h_t(\alpha)$ bởi $h_t(\hat{\alpha}_{pr})$, thì thu được

$$\frac{y_t}{h_t(\hat{\alpha}_{pr})} \approx \left(\frac{X_{t-1}}{h_t(\hat{\alpha}_{pr})}\right)^T .\alpha + \eta_t \tag{2.13}$$

Chúng ta lại bỏ qua tính ngẫu nhiên trong $h_t(\hat{\alpha}_{pr})$, thì xem đây là 1 mô hình hồi quy chuẩn với sai số có phương sai có điều kiện không đổi. Do đó, ước lượng bằng phương pháp bình phương cực tiểu thu được:

$$\hat{\alpha} = \left(\sum_{t=1}^{n} \frac{X_{t-1}X_{t-1}^T}{h_t^2(\hat{\alpha}_{pr})}\right)^{-1} . \left(\sum_{t=1}^{n} \frac{X_{t-1}Y}{h_t^2(\hat{\alpha}_{pr})}\right) \tag{2.14}$$

Trước hết, ta có kết quả về phân bố tiệm cận của $\hat{\alpha}_{pr}$.

Định lý 2.2 Trong mô hình (2.19) và (2.20), giả thiết $\forall\ 1 \le j, k, l, m \le Q$:

$$E\{y_j y_k y_l y_m\} < \infty \tag{2.15}$$

Khi đó $\hat{\alpha}_{pr}$ thoả mãn

$$n^{1/2}(\hat{\alpha}_{pr} - \alpha) \xrightarrow{D} N\left[0,\ Var(\upsilon_1^2)\{E(X_o X_o^T)\}^{-1} . E\{(\alpha^T X_o)^2 X_o X_o^T\}\{E(X_o X_o^T)\}^{-1}\right] \tag{2.16}$$

Chú ý rằng điều kiện (2.16) đảm bảo rằng $E(X_o X_o^T)$ và $E\{(\alpha^T X_o)^2 X_o X_o^T\}$ là xác định. Khi các sai số là chuẩn, điều kiện đủ cho sự tồn tại các mômen cấp cao của y_t trong biểu thức của α được cho bởi Engle (1982, Định lý 1 và 2, [19])

Từ kết quả của định lý này, ta chứng minh được kết quả về phân bố tiệm cận của $\hat{\alpha}$.

Định lý 2.3 Trong mô hình ARCH với giả thiết về tính chất của $\hat{\alpha}_{pr}$:

$$n^{1/2}(\hat{\alpha}_{pr} - \alpha) = O_p(1) \tag{2.17}$$

Và tính chất mômen của $\{\varepsilon_i; 1 - Q \le i \le 0\}$: $\forall\, 1 \le j, k, l \le Q$,

$$E\left\{y_{-j}y_{-k}y_{-l} / \left(\alpha^T X_o\right)^3\right\} < \infty \tag{2.18}$$

Khi đó, phân bố của $\hat{\alpha}$ có tính chất

$$n^{1/2}(\hat{\alpha} - \alpha) \xrightarrow{D} N\left[0,\ \mathrm{Var}(\upsilon_1^2)\left\{E\left[X_0 X_o^T \left(\alpha^T X_o\right)^2\right]\right\}^{-1}\right] \tag{2.19}$$

Chú ý rằng các điều kiện của định lý 2.2 suy ra (2.17). Đặc biệt, trong phân bố tiệm cận của ước lượng QMLE, Weiss (1986) giả sử $\forall\, 1 \le j, k \le Q$: $E\left\{y_{-j}y_{-k} / \left(\alpha^T X_o\right)^2\right\} < \infty$. Khi $\alpha_j > 0\ \forall\, 1 \le j \le Q$, điều kiện (2.18) hiển nhiên thoả mãn vì hàm $y \to y / (\alpha_o + \alpha_j y)$ là bị chặn.

Từ kết quả của định lý 2.3 và định lý 2.4, ta có các bước chính của thuật toán ước lượng tham số của mô hình ARCH(Q).

Thuật toán 2.3. Thuật toán ước lượng tham số của mô hình ARCH

Bước 1: Ước lượng bậc Q.

Viết mô hình (2.20) dưới dạng

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_Q \varepsilon_{t-Q}^2 + \varsigma_t \tag{2.20}$$

Ở đây, $\varsigma_t = \varepsilon_t^2 - h_t$ thoả mãn tính chất

$E(\varsigma_t) = 0$ và $E(\varsigma_t\, \varsigma_s) = 0\ (t \neq s;\ t, s = \overline{1, n})$

Khi đó, (2.20) có dạng một quá trình tự hồi quy AR(Q) với sai số $\{\varsigma_t\}$ không tương quan có trung bình không. Dùng các tiêu chuẩn AIC hoặc AICC hoặc dùng đồ thị hàm tự tương quan và tự tương quan riêng đối với quá trình tự hồi quy AR(Q) [16] có biểu diễn (2. 20) để ước lượng bậc Q của mô hình ARCH

Bước 2: Ước lượng $\hat{\alpha}_{pr}$ của α bằng phương pháp bình phương cực tiểu theo mô hình

$$y_t = X_{t-1}^T.\alpha + h_t(\alpha).\eta_t;\ t = \overline{1, n}$$

Bước 3: Ước lượng các phương sai có điều kiện

$$h_t(\hat{\alpha}_{pr}) = X_{t-1}^T\hat{\alpha}_{pr};\ t = \overline{1,n}$$

Bước 4: Ước lượng $\hat{\alpha}$ của α bằng phương pháp bình phương cực tiểu theo mô hình

$$\frac{y_t}{h_t(\hat{\alpha}_{pr})} \approx \left(\frac{X_{t-1}}{h_t(\hat{\alpha}_{pr})}\right)^T.\alpha + \eta_t;\ t = \overline{1,n}$$

2.3.5. Kiểm định sự phù hợp của mô hình ARCH(Q)

Trong định nghĩa quá trình ARCH(Q) luôn giả thiết $\upsilon_t = \varepsilon_t / \sqrt{h_t} \sim IID(0, 1)$. Khi đó, mô hình ARCH(Q) được gọi là phù hợp nếu $\upsilon_t \sim IID(0, 1)$. Do vậy, ta có các bước chính của thuật toán cài đặt kiểm định sự phù hợp của mô hình ARCH(Q).

Thuật toán 2.4. Thuật toán kiểm định sự phù hợp của mô hình ARCH(Q) trải qua các bước chính

Bước 1: Ước lượng các phương sai có điều kiện của mô hình ARCH

$$h_t = E(\varepsilon_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{Q}\alpha_i};\ t = \overline{1,Q}$$

$$h_t = \alpha_0 + \sum_{i=1}^{Q}\alpha_i\varepsilon_{t-i}^2;\ t = \overline{Q+1,n}$$

Bước 2: Ước lượng $\upsilon_t = \varepsilon_t / \sqrt{h_t}$; $t = \overline{1,n}$

Bước 3: Kiểm định $\upsilon_t \sim IID(0, 1)$

Mục đích của kiểm định $\{\upsilon_t\}$ - nhiễu chuẩn hoá là tính toán các tham số đặc trưng của nó so sánh với các đặc trưng của phân bố chuẩn. Từ nhiễu chuẩn hoá này, tính giá trị của hàm tự tương quan và biểu diễn trên đồ thị. Sau đó, thực hiện các kiểm định nhằm phát hiện sự biến mất của hiệu ứng ARCH.

Từ các kết quả phân tích trên, ta có các bước chính của thuật toán cài đặt nhận dạng mô hình ARCH(Q).

Giả sử chuỗi thời gian đầu vào $\mathscr{X} = \{x_t;\ t = \overline{1,n}\}$

Thuật toán 2.5. Thuật toán nhận dạng mô hình ARCH(Q) trải qua các bước chính

Bước 1: Ước lượng mô hình ARMA(p, q) tốt nhất cho $\mathscr{X}$.

$$X_t - a_1 X_{t-1} - \ldots\ldots - a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \ldots. + b_q \varepsilon_{t-q}$$

Bước 2: Ước lượng nhiễu thu được từ mô hình ARMA(p, q), bình phương nhiễu thu được $\varepsilon_t^2;\ t = \overline{1,n}$. Tiến hành kiểm định hiệu ứng ARCH.

* Nếu không có hiệu ứng ARCH: dừng thuật toán, kết luận chuỗi quan sát $\mathscr{X}$ thoả mãn mô hình ARMA(p, q).

* Nếu có hiệu ứng ARCH: chuyển sang bước 3.

Bước 3: Ước lượng tham số của mô hình ARCH(Q).

Bước 4: Kiểm định sự phù hợp của mô hình ARCH(Q).

2.4. Qúa trình tự hồi quy với những biến động bất thường tổng quát – GARCH (Generalized AutoRegresive Conditional Heteroskedasticity)

2.4.1. Khái niệm

Giả sử $\{\varepsilon_t\}$ là quá trình ngẫu nhiên nhận giá trị rời rạc ($\varepsilon_t = r_t - \mu_t : r_t = \log P_t; \mu_t$ là trung bình của r_t) và F_t là tập thông tin quá khứ (σ - trường) tạo bởi các giá trị ε_t, ε_{t-1}, ... Khi đó quá trình GARCH(m, s) được định nghĩa bởi

$$\varepsilon_t\big|_{F_{t-1}} \sim \mathrm{IID}(0, h_t) \Leftrightarrow \varepsilon_t = \upsilon_t.\sigma_t;\ \text{với } \{\upsilon_t\} \sim \mathrm{IID}(0,1)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

Ở đây m, s $\in N$

$$\alpha_0 > 0;\ \alpha_i \geq 0, \beta_j \geq 0 \text{ và } \sum_{i=1}^{\max\{m,s\}} (\alpha_i + \beta_i) < 1$$

Trong thực hành, thường giả thiết $\{\upsilon_t\} \sim N(0,1)$. Nếu s = 0 thì qúa trình GARCH suy biến thành quá trình ARCH(m).

2.4.2. Phương pháp ước lượng tham số

Bằng cách đặt ẩn phụ $\upsilon_t = \varepsilon_t^2 - \sigma_t^2$, khi đó mô hình GARCH trở thành

$$\varepsilon_t^2 = \alpha_o + \sum_{i=1}^{m} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{s} \beta_j \varepsilon_{t-j}^2 + \upsilon + \sum_{j=1}^{s} (-\beta_j) \upsilon_{t-j}$$

Đây là mô hình có dạng ARMA(max(r, s); s) đối với ε_t^2. Sử dụng phương pháp ước lượng của mô hình ARMA ta xác định được các tham số của mô hình GARCH. Tuy nhiên với giả thiết $\{\upsilon_t\}$ không còn là IID(0, 1).

2.5. Quá trình GARCH kết hợp (IGARCH)

Trong trường hợp đa thức của biểu diễn GARCH trong mục 2.4 có nghiệm đơn vị thì ta thu được mô hình IGARCH. Do đó, các mô hình IGARCH là các mô hình GARCH có nghiệm đơn vị. Chẳng hạn một quá trình IGARCH(1; 1) có thể được viết bởi: $\varepsilon_t = \upsilon_t.\sigma_t$; với $\{\upsilon_t\} \sim \text{IID}(0,1)$

$$\sigma_t^2 = \alpha_o + \beta_1 \sigma_{t-1}^2 + (1-\beta_1)\varepsilon_{t-1}^2 \text{ ;trong đó } 0 < \beta_1 < 1$$

Phương pháp ước lượng tham số sử dụng các phương pháp của quá trình ARIMA.

Các thuật toán nhận dạng mô hình GARCH, IGARCH ta xây dựng tương tự như mô hình ARCH.

CHƯƠNG 3. DỰ BÁO BIẾN ĐỘNG GIÁ CHỨNG KHOÁN VÀ ÁP DỤNG VÀO THỊ TRƯỜNG CHỨNG KHOÁN VIỆT NAM

Nội dung của chương 3 của đề tài, ứng dụng các mô hình chuỗi thời gian ARCH/ GARCH vào bài toán dự báo biến động giá chứng khoán và áp dụng vào thị trường chứng khoán Việt Nam. Chương này chủ yếu nghiên cứu 2 vấn đề:

- Xây dựng công thức dự báo cho các quá trình dừng, đặc biệt là quá trình ARCH/ GARCH. Sử dụng các kết quả của việc nhận dạng mô hình và kiểm định sự phù hợp của mô hình để xây dựng các dự báo điểm cho chuỗi thời gian. Từ việc xây dựng các dự báo điểm đối với quá trình ARCH/ GARCH, ta sẽ chỉ ra được trong trường hợp, phương sai có điều kiện của nhiễu $\{\varepsilon_t\}$ khác hằng số thì ảnh hưởng như thế nào đến độ chính xác của dự báo.

- Áp dụng các kết quả lý thuyết đối với các chuỗi số liệu cụ thể thuộc 2 lĩnh vực:

+) Một số bộ số liệu về chỉ số chứng khoán của thị trường chứng khoán thế giới

+) Một số bộ số liệu về chỉ số chứng khoán của thị trường chứng khoán Việt Nam.

3.1. Dự báo tuyến tính theo các mô hình chuỗi thời gian

Các thuật toán dùng để tính toán giá trị dự báo của quá trình ngẫu nhiên dừng $\{X_t\}$ như thuật toán Durbin – Lewinson, thuật toán Burg, thuật toán đổi mới [17] cho phép tính toán trực tiếp giá trị dự báo $\hat{X}_{n+h}$ của X_{n+h} trên cơ sở $\mathcal{X} = \{x_t;\ t = \overline{1,n}\}$ có thể tham khảo tại nhiều tài liệu [1, 17] và cũng được cài đặt trong nhiều phần mềm thống kê như Eviews hoặc Pest.

Các phương pháp tính toán giá trị dự báo của chuỗi thời gian dừng (hoặc trong trường hợp tổng quát là quá trình có mômen cấp 2 hữu hạn) sử dụng trong đề tài đều áp dụng các kết quả nhận dạng mô hình ARMA có hiệu ứng ARCH, mô hình ARCH/ GARCH sau khi đã kiểm tra sự phù hợp của mô hình chuỗi thời gian.

3.1.1. Dự báo tuyến tính theo mô hình ARMA với hiệu ứng ARCH

Trong phương pháp dự báo sử dụng mô hình ARCH/ GARCH giới thiệu ở chương 2 luôn giả thiết $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ hoặc $\{\varepsilon_t\} \sim IID(0, \sigma^2)$ và phương sai có điều kiện của $\{\varepsilon_t\}$ đối với tập thông tin F_{t-1} là hằng số (Ở đây, F_t là σ - trường chứa x_t, ε_t, x_{t-1}, ε_{t-1}, ...). Trong trường hợp ε_t không là ồn trắng, mà ε_t tuân theo mô hình ARCH(Q) thì tập thông tin F_t sẽ ảnh hưởng tới độ chính xác của dự báo. Trong phần này ta sẽ nghiên cứu ảnh hưởng đó tới sai số dự báo của mô hình ARMA.

Giả sử từ chuỗi quan sát $\mathscr{X} = \{x_t;\, t = \overline{1,n}\}$, nhận dạng mô hình ARMA(p, q):

$$a(B)X_t = b(B)\varepsilon_t \tag{3.1}$$

$$\varepsilon_t \sim ARCH(Q)$$

Hay $\varepsilon_t\big|_{F_{t-1}} \sim IID(0, h_t) \Leftrightarrow \varepsilon_t = \upsilon_t\sqrt{h_t};\, \upsilon_t \sim IID(0, 1)$ (3.2)

$$h_t = \alpha_0 + \sum_{i=1}^{Q}\alpha_i\varepsilon_{t-i}^2 \tag{3.3}$$

Với điều kiện $Q > 0;\ \alpha_0 > 0;\ \alpha_i \geq 0,\ i = \overline{1,Q};\ \sum_{i=1}^{Q}\alpha_i < 1$

Hiệu ứng ARCH xuất hiện khi có sự tồn tại của mômen cấp cao của ε_t^2 nên dự báo điểm $f_{n,h}$ theo mô hình (3.1) cũng chính là dự báo điểm theo mô hình (3.1). Trong khi đó, phương sai sai số dự báo

$$Var(e_{n,h}\big|_{F_n}) = \sum_{i=0}^{h-1}\varphi_i^2 E\left(\varepsilon_{n+h-i}^2\big|_{F_n}\right) \tag{3.4}$$

Khi xuất hiện hiệu ứng ARCH, $E(\varepsilon_{n+h-i}^2\big|_{F_n})$ sẽ phụ thuộc vào các phần tử của F_n và do đó sẽ phụ thuộc vào n. Trong trường hợp ngược lại, $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ hoặc $\{\varepsilon_t\} \sim IID(0, \sigma^2)$, tức là mô hình có phương sai có điều kiện hằng số thì $E(\varepsilon_{n+h-i}^2\big|_{F_n}) = \sigma^2$ nên phương sai của sai số dự báo

$$Var(e_{n,h}\big|_{F_n}) = \sigma^2\sum_{i=0}^{h-1}\varphi_i^2 \tag{3.5}$$

Trong trường hợp này, phương sai của sai số dự báo chỉ phụ thuộc vào h (độ dài của thời kỳ dự báo) chứ không phụ thuộc vào tập thông tin F_n. Sử dụng (3.4) để xây dựng khoảng dự báo cần phải tính $E(\varepsilon^2_{n+h-i}|_{F_n})$. Điều này có thể làm được bằng cách biểu diễn ARCH như là một quá trình AR(Q).

$$\varepsilon_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + ... + \alpha_Q\varepsilon_{t-Q}^2 + \varsigma_t \qquad (3.6)$$

Ở đây, $\varsigma_t = \varepsilon_t^2 - h_t$ thoả mãn tính chất $E(\varsigma_t) = 0$ và $E(\varsigma_t\varsigma_s) = 0;\ t \neq s$.

Từ (3.6), ta có

$$\varepsilon^2_{n+h} = \alpha_0 + \sum_{i=1}^{Q}\alpha_1\varepsilon^2_{n+h-i} + \varsigma_{n+h}$$

Lấy kỳ vọng có điều kiện 2 vế ta được:

$$E(\varepsilon^2_{n+h}|_{F_n}) = \alpha_0 + \sum_{i=1}^{Q}\alpha_1 E(\varepsilon^2_{n+h-i}|_{F_n}) + E(\varsigma_{n+h}|_{F_n}) \qquad (3.7)$$

Ở đây, ta có

i) $E(\varepsilon^2_{n+h-i}|_{F_n}) = \varepsilon^2_{n+h-i}$ nếu $i \geq h$;

ii) $E(\varepsilon^2_{n+h-i}|_{F_n})$ với $i < h$, tính đệ quy theo (3.7);

iii) $E(\varsigma_{n+h}|_{F_n}) = 0$.

Để định ý, xét quá trình AR(1)

$$x_t = \phi_1 x_{t-1} + \varepsilon_t; |\phi_1| < 1$$

$$\varepsilon_t \sim ARCH(1) \Leftrightarrow \begin{cases} \varepsilon_t = \upsilon_t.\sqrt{h_t} \\ h_t = \alpha_0 + \alpha_1\varepsilon^2_{t-1} \end{cases}$$

Dự báo điểm bước h xác định bởi công thức:

$$f_{n,h} = \phi_1 f_{n,h-1};\ h = 1,2,3,...$$

Phương sai sai số dự báo bước h:

$$Var(e_{n,h}|_{F_n}) = \sum_{i=0}^{h-1}\phi_1^{2i}E(\varepsilon^2_{n+h-i}|_{F_n});\ h = 1, 2, 3, ... \qquad (3.8)$$

Sử dụng biểu diễn AR(1) cho ARCH(1):

$$\varepsilon^2_{n+h} = \alpha_0 + \alpha_1\varepsilon^2_{n+h-1} + \varsigma_{n+h}$$

Nên $E(\varepsilon^2_{n+h}|_{F_n}) = \alpha_0 + \alpha_1 E(\varepsilon^2_{n+h-1}|_{F_n})$ (3.9) với $E(\varepsilon^2_{n+h-i}\big|_{F_n}) = \varepsilon^2_{n+h-i}$ nếu $i \geq h$. (3.10)

Khi đó, phương sai của sai số dự báo (3.8) tính toán sử dụng (3.9) và (3.10).
Từ đây, thu được thuật toán dự báo theo mô hình ARMA có hiệu ứng ARCH.

3.1.2. Dự báo tuyến tính theo mô hình ARCH

Giả sử quá trình ngẫu nhiên $\varepsilon_t \sim ARCH(Q)$, tức là

$$\varepsilon_t\big|_{F_{t-1}} \sim IID(0, h_t) \Leftrightarrow \varepsilon_t = \upsilon_t.\sqrt{h_t}\text{, với } \upsilon_t \sim IID(0,1)$$

$$h_t = \alpha_o + \sum_{i=1}^{Q}\alpha_i\varepsilon^2_{t-i} \quad (3.11)$$

Ở đây $Q > 0$; $\alpha_0 > 0$; $\alpha_i \geq 0, i = \overline{1,Q}$; $\sum_{i=1}^{Q}\alpha_i < 1$

Tập thông tin F_t là σ - trường sinh bởi $\{\varepsilon_t, \varepsilon_{t-1}, ...\}$

Như đã chứng minh ở 2.3.2, $Var(\varepsilon_t / F_{t-1}) = h_t = \alpha_o + \sum_{i=1}^{Q}\alpha_i\varepsilon^2_{t-i}$.

Phương sai có điều kiện h_t đo mức độ phân tán hay độ biến động của ε_t quanh giá trị trung bình. Nếu chúng ta dự đoán được sự thay đổi của h_t trên cơ sở tập thông tin F_{t-1} là có ý nghĩa rất lớn trong thực tế. Đặc biệt trong lĩnh vực tài chính, nếu ta biết được sự biến động lên xuống của giá của 1 cổ phiếu nào đó vào ngày mai thì sẽ có lợi nhiều trong việc quyết định một phương án đầu tư không những mang lại hiệu quả kinh tế cao mà còn tránh được các rủi ro. Do đó, ta đặt vấn đề tính giá trị h_t tại thời điểm n + k trên cơ sở tập quan sát $\mathscr{X} = \{\varepsilon_t; t = \overline{1,n}\}$.

Dự báo theo mô hình ARCH(Q) cũng có thể xây dựng tương tự như dự báo theo mô hình AR(Q).

Thật vậy, gọi $f_{n,1}$ là dự báo bước 1 của h_{n+1} trên cơ sở F_n. Khi đó, ta có

$$f_{n,1} = E(h_{n+1}\big|_{F_n}) = E\left(\alpha_o + \sum_{i=1}^{Q}\alpha_i\varepsilon^2_{n+1-i}\bigg|_{F_n}\right)$$

$$= \alpha_o + \sum_{i=1}^{Q}\alpha_i E\left(\varepsilon^2_{n+1-i}\big|_{F_n}\right)$$

$$= \alpha_o + \sum_{i=1}^{Q} \alpha_i \varepsilon_{t-i}^2$$

Tương tự, dự báo k bước của h_{n+1} trên cơ sở F_n là

$$f_{n,k} = E(h_{n+k}|_{F_n}) = E\left(\alpha_o + \sum_{i=1}^{Q} \alpha_i \varepsilon_{n+k-i}^2 \Big|_{F_n} \right)$$

$$= \alpha_o + \sum_{i=1}^{Q} \alpha_i f_{n,k-i}$$

Ở đây, $f_{n,k-i} = \varepsilon_{n+k-i}^2$ nếu $k - i \le 0$.

Từ đây, thu được thuật toán dự báo theo mô hình ARCH.

3.1.3. Dự báo tuyến tính theo mô hình GARCH

Để đơn giản, ta xét mô hình GARCH(1; 1) với

$$\sigma_t^2 = \alpha_o + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, 0 \le \alpha_1, \beta_1 \le 1, (\alpha_1 + \beta_1) < 1 \tag{3.12}$$

Từ (3.12) cần xây dựng công thức dự báo k bước với gốc thời gian h

* Dự báo bước 1: $\sigma_h^2(1) = \sigma_{h+1}^2 = \alpha_o + \alpha_1 \varepsilon_h^2 + \beta_1 \sigma_h^2$

* Để xác định dự báo bước k: viết lại phương trình biến động

$$\sigma_{t+1}^2 = \alpha_o + (\alpha_1 + \beta_1)\sigma_t^2 + \alpha_1 \sigma_t^2 (\varepsilon_t^2 - 1)$$

Khi t = h + 1, ta có $\sigma_{t+2}^2 = \alpha_o + (\alpha_1 + \beta_1)\sigma_{h+1}^2 + \alpha_1 \sigma_{h+1}^2 (\varepsilon_{h+1}^2 - 1)$. Suy ra dự báo bước 2 là $\sigma_h^2(2) = \alpha_o + (\alpha_1 + \beta_1)\sigma_h^2(1)$

Tổng quát, ta có công thức dự báo bước h là:

$$\sigma_h^2(k) = \alpha_o + (\alpha_1 + \beta_1)\sigma_h^2(k-1);\ k > 1 \tag{3.13}$$

Đây cũng tương tự như công thức dự báo cho mô hình ARMA(1; 1) với đa thức tự hồi quy $1 - (\alpha_1 + \beta_1)B$. Bằng cách lặp lại công thức (3.13) sau k lần ta thu được

$$\sigma_h^2(k) = \frac{\alpha_o \left[1 - (\alpha_1 + \beta_1)^{k-1}\right]}{1 - (\alpha_1 + \beta_1)} + (\alpha_1 + \beta_1)^{k-1} \sigma_h^2(1)$$

Do đó $\sigma_h^2(k) \to \dfrac{\alpha_o}{1 - \alpha_1 - \beta_1}$ khi $k \to \infty$.

3.1.4. Dự báo tuyến tính theo mô hình IGARCH

Để đơn giản, ta xét mô hình IGARCH(1; 1) với

$$\sigma_t^2 = \alpha_o + \beta_1\sigma_{t-1}^2 + (1-\beta_1)\varepsilon_{t-1}^2 \quad (3.14)$$

Từ (3.14) cần xây dựng công thức dự báo k bước với gốc thời gian h

* Dự báo bước 1: $\sigma_h^2(1) = \alpha_o + \beta_1\sigma_h^2 + (1-\beta_1)\varepsilon_h^2$

Tổng quát, ta có công thức dự báo bước k là:

$$\sigma_h^2(k) = \sigma_h^2(1) + (k-1)\alpha_o;\quad k \geq 1 \quad (3.15)$$

Từ công thức (3.15) ta thấy phương trình dự báo theo mô hình IGARH(1, 1) tuân theo một đường thẳng có hệ số góc α_o.

3.2. Ứng dụng các mô hình chuỗi thời gian tài chính vào bài toán dự báo dự báo biến động giá chứng khoán

Các kết quả số trong đề tài đều được thực hiện cùng phần mềm *Phân tích chuỗi thời gian* đi kèm cùng với đề tài. Phần mềm này ứng dụng tất cả các thuật toán đã được đề cập trong đề tài và một số thuật toán khác. Phần mềm P*hân tích chuỗi thời gian* được thực hiện dựa trên ngôn ngữ lập trình Visual Basic 6. 0 và có thể chạy trên mọi hệ thống máy tính tương thích chuẩn x86 – 32 bit hoặc x86 – 64 bit sử dụng các phiên bản 32 bit hoặc 64 bit của hệ điều hành Microsoft Windows.

3.2.1. Biến động - đặc trưng của biến động

Biến động là một nhân tố quan trọng trong các giao dịch hợp đồng quyền lựa chọn. Ở đây biến động được hiểu là phương sai có điều kiện của lãi suất tài sản cơ sở. Chẳng hạn, giá của một quyền chọn kiểu châu Âu, là hợp đồng mà cho phép người giữ nó một quyền thực thi, nhưng không phải là nghĩa vụ là mua một số cổ phiếu của một chứng khoán định rõ với một mức giá cố định vào một ngày cho trước. Giá cố định đó gọi là *giá thực thi*, ký hiệu là K, ngày cho trước gọi là *ngày đáo hạn*, khoảng thời gian đáo hạn là ℓ. Trong trường hợp người giữ có thể thực thi quyền chọn tại bất kỳ thời điểm nào trước ngày đáo hạn thì quyền chọn đó được gọi là quyền chọn kiểu Mỹ. Khi đó, xét công thức Black – Scholes về giá của quyền chọn mua kiểu châu Âu xác định bởi:

$$c_t = P_t\Phi(x) - K.r^{-\ell}.\Phi(x - \sigma_t.\sqrt{\ell}) \quad (3.16)$$

Và $$x = \frac{\ln(P_t / Kr^{-\ell})}{\sigma_t\sqrt{\ell}} + \frac{1}{2}\sigma_t\sqrt{\ell} \qquad (3.17)$$

Trong đó P_t giá cổ phiếu tại thời điểm t, r là lãi suất không rủi ro, σ_t là độ lệch chuẩn có điều kiện của log lợi nhuận của chứng khoán cụ thể và $\Phi(x)$ là hàm phân bố xác suất của biến ngẫu nhiên phân bố chuẩn. Công thức này có nhiều phiên bản khác nhau và ngày nay được mở rộng theo nhiều hướng khác nhau nhằm phản ánh chính xác hơn những diễn biến của thị trường chứng khoán. Tuy nhiên ở đây, ta thấy vai trò của σ_t trong công thức (3.16), biến động σ_t thay đổi theo thời gian.

Biến động có ý nghĩa rất quan trọng trong quản lý rủi ro, mô hình biến động cung cấp một nhánh tiếp cận đơn giản để tính toán giá trị rủi ro của tình trạng tài chính. Mặt khác mô hình biến động trong chuỗi thời gian có thể đóng vai trò quan trọng trong việc cải thiện hiệu quả của ước lượng tham số và đánh giá độ chính xác của dự báo.

Đặc trưng của biến động: Một đặc trưng cơ bản của biến động của chứng khoán thường không quan sát trực tiếp được, chẳng hạn chúng ta xét log lãi suất theo ngày của chứng khoán SAM (Đồ thị hình 3.10): biến động theo ngày không quan sát được từ lãi suất bởi vì chỉ có một quan sát trong một ngày giao dịch. Nếu dữ liệu trong ngày của chứng khoán chẳng hạn như lãi suất từng giờ có hiệu lực thì có thể ước lượng được biến động theo ngày. Hơn nữa biến động của chứng khoán có thể chứa biến động ngày và sự thay đổi giữa các ngày giao dịch. Khả năng không quan sát được của biến động cũng ảnh hưởng tới khả năng thực hiện dự báo theo mô hình ARCH(Q) hoặc GARCH(P, Q).

Trong thị trường quyền chọn mua, nếu chấp nhận phương án giá bị chi phối bởi một mô hình kinh tế như mô hình Black – Scholes, thì có thể thu được biến động từ công thức về giá. Một hướng tiếp cận là dùng các mô hình chỉ định, mà dựa trên các giả thiết có thể không cố định trong thực hành. Chẳng hạn, từ giá quan sát được của hợp đồng quyền chọn châu Âu ta có thể ước lượng độ lệch có điều kiện σ_t từ công thức Black – Scholes, giá trị thu được σ_t^2 gọi là biến động suy ra từ các chứng khoán cơ sở. Tuy nhiên, để tính toán được biến động này, ta sử dụng giả thiết log phân bố chuẩn cho chuỗi lãi suất. Điều này rất khó có thể xảy ra trong thực tế. Kinh nghiệm cho thấy biến động thu được theo cách này, thường lớn hơn so với biến động thu được từ mô hình ARCH(Q) hoặc GARCH(P,Q) [22].

Mặc dầu biến động không quan sát trực tiếp được, nhưng đối với lãi suất của các tài sản cơ sở, chúng có một vài đặc trưng tổng quát [34, 36, 37]:

- Luôn tồn tại các cụm biến động.
- Biến động thay đổi theo thời gian.
- Biến động không hội tụ tới ∞, tức là biến động biến thiên trong phạm vi cố định. Hay nói cách khác biến động thường là thể hiện của quá trình dừng.
- Biến động dường như cũng ảnh hưởng trở lại tới sự tăng giá mạnh hoặc giảm giá mạnh.

Những tính chất này của biến động có tác dụng rất quan trọng trong việc ứng dụng các mô hình biến động trong thực hành.

3.2.2. Cấu trúc của mô hình biến động

Gọi $r_t = \ln(P_t) - \ln(P_{t-1})$, P_t là giá của chứng khoán tại thời điểm t. Ý tưởng cơ bản để nghiên cứu biến động là xét xem chuỗi $\{r_t\}$ hoặc không có tương quan thứ tự hoặc có tương quan thứ tự cấp thấp nhưng phụ thuộc.

Để nghiên cứu các tính chất của biến động, cần xem xét kỳ vọng có điều kiện và phương sai có điều kiện của r_t đối F_{t-1} là σ - trường thông tin tại thời điểm t - 1 là

$$\mu_t = E(r_t / F_{t-1}); \; \sigma_t^2 = Var(r_t / F_{t-1}) = E\left[(r_t - \mu_t)^2 / F_{t-1}\right] \qquad (3.18)$$

Mối liên hệ tương quan có thứ tự của chuỗi $\{r_t\}$ có thể mô hình hoá bởi một mô hình chuỗi thời gian ARMA(p, q) cho $\{r_t\}$ như sau:

$$r_t = \mu_t + \varepsilon_t; \; \mu_t = a_0 + \sum_{i=1}^{p} a_i r_{t-i} + \sum_{j=1}^{q} b_j \varepsilon_{t-j} \qquad (3.19)$$

Trong đó, p và p là các số nguyên không âm.

Mô hình (3.19) là một hướng ứng dụng của chuỗi thời gian tuyến tính (mô hình ARMA(p, q)) trong tài chính. Cấp của mô hình ARMA(p, q) phụ thuộc vào tần suất của chuỗi lãi suất, được xác định theo tiêu chuẩn AIC, AICC [17]. Chẳng hạn, lãi suất hàng ngày của chỉ số thị trường thường diễn tả tương quan thứ tự thấp, nhưng lãi suất theo tháng không chứa tương quan thứ tự. Một cách khác để có thể nhận dạng mô hình cho μ_t là dùng biến giả để nghiên cứu ảnh hưởng của một ngày trong tuần đối với chuỗi lãi suất chứng khoán theo ngày.

Kết hợp phương trình (3.18) và (3.19), ta có

$$\sigma_t^2 = \mathrm{var}(r_t / F_{t-1}) = \mathrm{Var}(\varepsilon_t / F_{t-1}) \qquad (3.20)$$

Mô hình biến động bất thường cho { σ_t^2 } có thể phân loại theo 2 phạm trù:

- Mô hình mà dùng một hàm chính xác điều khiển sự tiến triển của σ_t^2.

- Mô hình dùng các phương trình ngẫu nhiên để mô tả σ_t^2.

Các mô hình ARCH hoặc GARCH thuộc phạm trù thứ nhất, và mô hình biến động ngẫu nhiên thuộc phạm trù thứ 2. đề tài chỉ nghiên cứu dạng thứ nhất. Các mô hình nghiên cứu trong đề tài: mô hình (3.43) đối với μ_t được gọi là *mô hình trung bình* cho r_t; còn mô hình cho σ_t^2 gọi là *phương trình biến động* cho r_t. Một trong những mô hình đó là mô hình ARCH(Q) của Engle (1982, [19]), Mô hình GARCH(r, s); Mô hình IGARCH(r,s), ….

Các tính chất của mô hình cũng như phương pháp và thuật toán nhận dạng mô hình đó đã giới thiệu ở chương 2 . Trong chương này ta sẽ ứng dụng các mô hình này vào bài toán dự báo biến động.

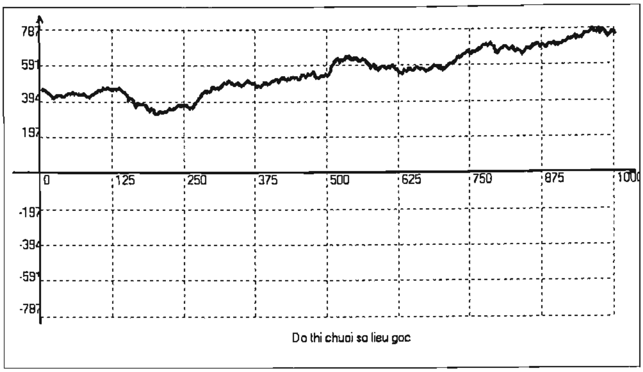
3.2.3. Nhận dạng mô hình ARCH(Q) đối với chuỗi thời gian tài chính

Các kết quả áp dụng mô hình ARCH(Q) đều thực hiện trên một số chuỗi thời gian tài chính kinh điển của thị trường tài chính thế giới và các chuỗi thời gian tài chính của Việt Nam (*Với lưu ý là: nguồn của các kết quả về đồ thị và kết quả của các mô hình cho ở trong đề tài này dựa trên nguồn gốc của số liệu đã trích dẫn và kết quả chạy phần mềm đi kèm với đề tài*)

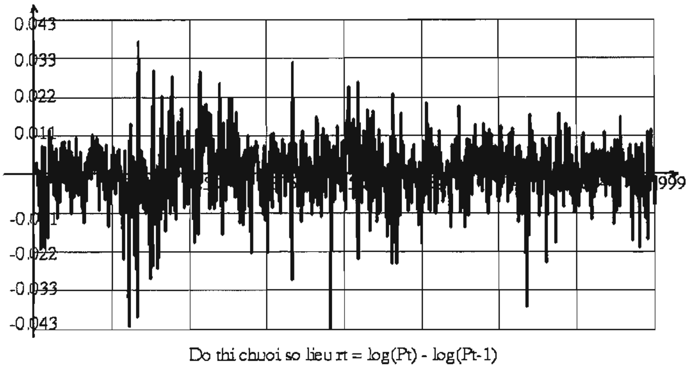
Mô hình 1. Chuỗi Nasdaq [38], Composite Stock Index Day Value

Số số liệu: 1000

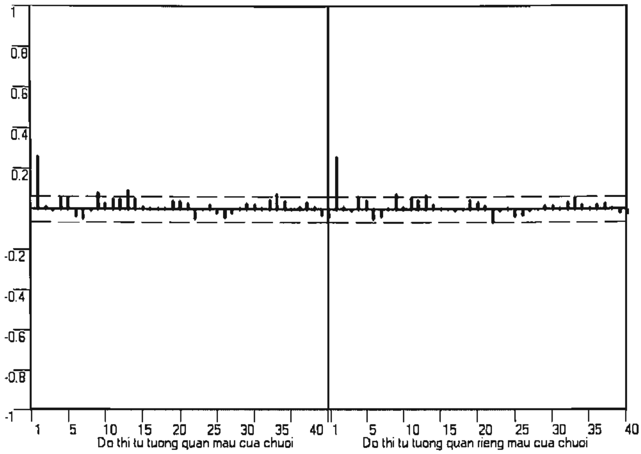
Nguồn gốc số liệu: [15]



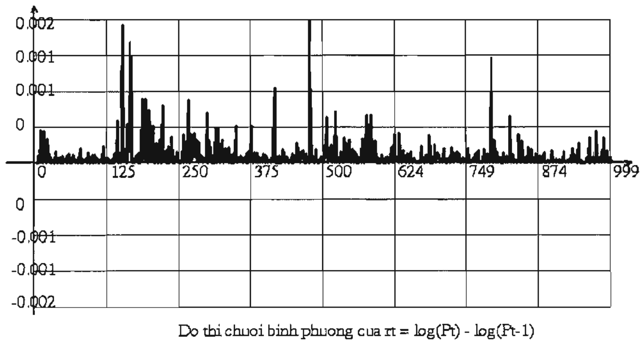
Hình 3.1. Đồ thị chuỗi chỉ số chứng khoán Nasdaq



Hình 3.2. Đồ thị chuỗi lãi suất chứng khoán Nasdaq

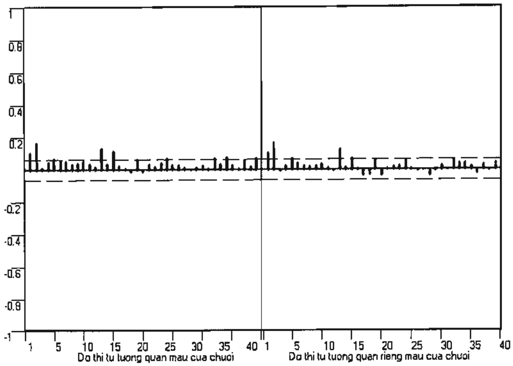


Hình 3.3. Đồ thị tự tương quan và tự tương quan riêng của chuỗi lãi suất.



Hình 3.4. Đồ thị chuỗi bình phương lãi suất chứng khoán Nasdaq

Đồ thị tự tương quan và tự tương quan riêng của chuỗi lãi suất bình phương.



Hình 3.5. Đồ thị tự tương quan và tự tương quan riêng chuỗi bình phương lãi suất chứng khoán.

Hiện tượng tạo cụm biến động và tự tương quan của chuỗi lãi suất và đặc biệt hiện tượng tự tương quan mạnh ở chuỗi bình phương lãi suất cho thấy có hiệu ứng ARCH đối với chuỗi lãi suất.

Ước lượng tham số bằng mô hình:

+) Ước lượng tham số bằng mô hình AR(1, 1) cho chuỗi $\{r_t\}_t$

$$r_t - 0{,}059328\, r_{t-1} = 0{,}00047\,4 + \varepsilon_t + 0{,}20816\, \varepsilon_{t-1}$$

+) Dùng thống kê Ljung – Box kiểm định hiệu ứng ARCH trên chuỗi $\{\varepsilon_t\}_t$

h	Q_{LB}	$\chi^2(h), \alpha = 0{,}05$
1	29,283	3,841
2	78,778	5,991
3	79,723	7,815
4	83,547	9,488
5	89,030	11,071
6	93,338	12,592
7	93,686	14,067
8	96,199	15,507
9	98,471	16,919
10	98,973	18,307

Bảng 3.6. Giá trị thống kê Q_{LB} và giá trị tới hạn $\chi^2(h)$

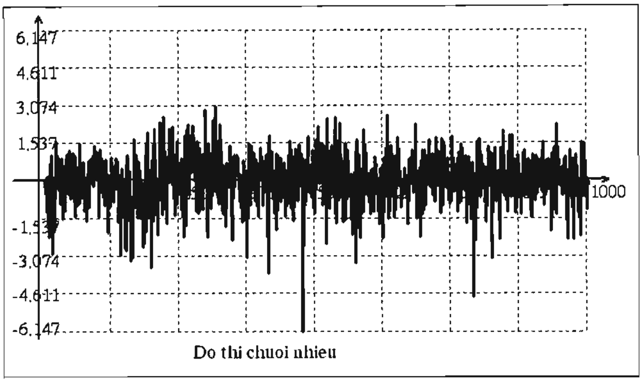
So sánh giá trị thống kê Q_{LB} với giá trị tới hạn khi bình phương ở bảng 3.7. Với mức ý nghĩa $\alpha = 0{,}05$; kết luận bác bỏ giả thuyết H_o và chấp nhận giả thuyết H_1 có hiệu ứng ARCH trên chuỗi lãi suất. Kết quả này phù hợp với sự xuất hiện hiệu ứng ARCH trên đồ thị chuỗi số liệu và đồ thị hàm tự tương quan.

+) Nhận dạng mô hình ARCH(2).

$$\varepsilon_t = \sqrt{h_t}.\upsilon_t;\ \upsilon_t \sim \text{IID}(0;\ 1)$$

$$h_t = 0{,}000048 + 0{,}13719\varepsilon_{t-1}^2 + 0{,}199087\varepsilon_{t-2}^2 \qquad (3.21)$$

Đồ thị nhiễu chuẩn hoá.



Hình 3.7. Đồ thị nhiễu $\{\upsilon_t\}$ ước lượng được từ mô hình (3.21)

+) Kiểm định sự phù hợp mô hình. Thống kê Q_{LB} cho chuỗi chuẩn hoá $\{\upsilon_t\}$.

h	Q_{LB}	$\chi^2(h), \alpha = 0{,}05$
1	2,248	3,841
2	2,378	5,991
3	2,444	7,815
4	2,320	9,488
5	4,491	11,071
6	4,942	12,592
7	5,168	14,067
8	5,168	15,507
9	6,499	16,919
10	6,509	18,307

Bảng 3.8. Giá trị thống kê Q_{LB} và giá trị tới hạn $\chi^2(h)$

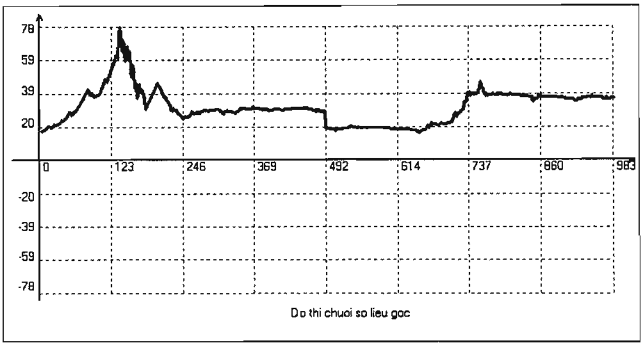
So sánh giá trị thống kê Q_{LB} với giá trị tới hạn khi bình phương ở bảng 3.8. Với mức ý nghĩa $\alpha = 0,05$; kết luận mô hình (3.21) là phù hợp.

➤ ***Kết luận 1.*** Xuất hiện hiệu ứng ARCH trên chuỗi Nasdaq, từ đó ta nhận dạng mô hình cho biến động h_t thu được ở (3.21). Các kết quả đối với các chuỗi thời gian tài chính kinh điển của thị trường chứng khoán thế giới, nhóm tác giả đề tài cũng thu được tương tự: đều xuất hiện hiệu ứng ARCH.

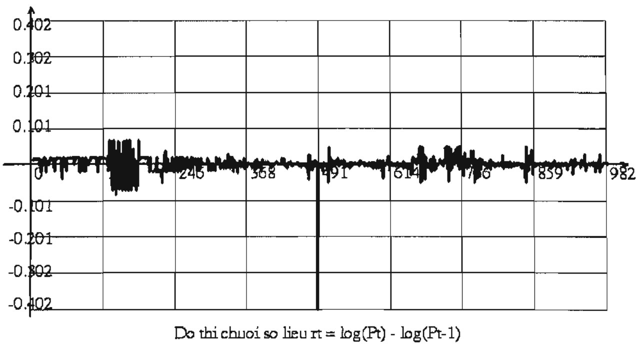
Mô hình 2. Chuỗi SAM [39] (chỉ số chứng khoán của công ty Cổ phần cáp và vật liệu viễn thông Sài Gòn).

Số số liệu: 983 (Thời điểm: 7/2000 đến 2004)

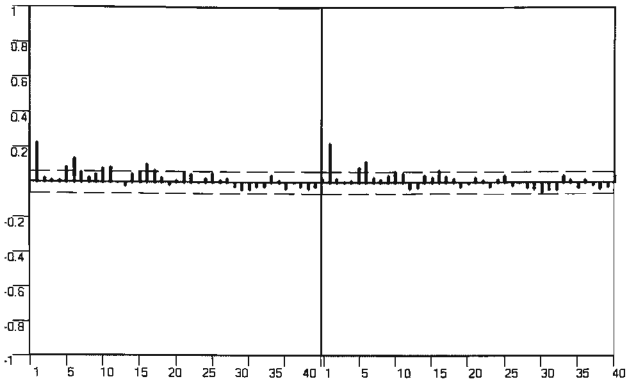
Nguồn số liệu:*http://www.hse.org.vn*



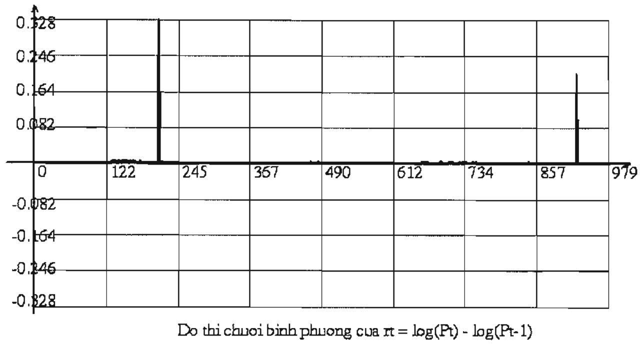
Hình 3.9. Đồ thị chuỗi chỉ số chứng khoán SAM



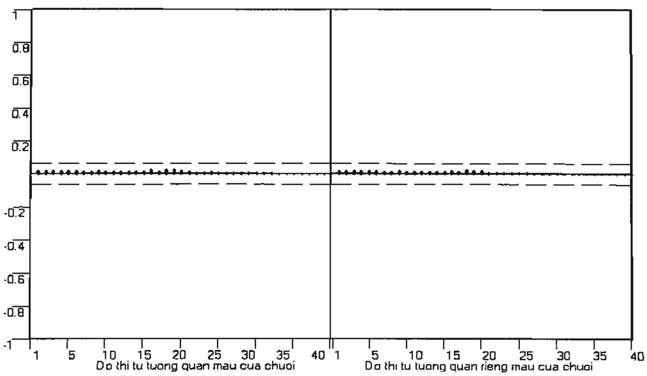
Hình 3.10. Đồ thị chuỗi lãi suất chứng khoán SAM



Hình 3.11. Đồ thị tự tương quan và tự tương quan riêng của chuỗi lãi suất.



Hình 3.12. Đồ thị chuỗi bình phương lãi suất chứng khoán.



Hình 3.13. Đồ thị tự tương quan và tự tương quan riêng chuỗi bình phương lãi suất chứng khoán.

Hiện tượng tạo cụm biến động ít và tính ít tương quan của chuỗi bình phương lãi suất chứng khoán cho thấy có thể không có hiệu ứng ARCH đối với chuỗi lãi suất. Ước lượng tham số bằng mô hình:

+) Ước lượng tham số bằng mô hình ARMA(1, 1) cho chuỗi $\{r_t\}_t$

$$r_t - 0{,}102068\, r_{t-1} = 0{,}000709 + \varepsilon_t + 0{,}121593\varepsilon_{t-1}$$

+) Dùng thống kê Ljung – Box kiểm định hiệu ứng ARCH trên chuỗi $\{\varepsilon_t\}_t$

h	Q_{LB}	$\chi^2(h), \alpha = 0{,}05$
1	0,598	3,841
2	0,742	5,991
3	0,903	7,815
4	1,061	9,488
5	1,19	11,071
6	1,369	12,592
7	1,459	14,067
8	1,528	15,507
9	1,717	16,919
10	1,83	18,307

Bảng 3.14. Giá trị thống kê Q_{LB} và giá trị tới hạn $\chi^2(h)$

So sánh giá trị thống kê Q_{LB} với giá trị tới hạn khi bình phương ở bảng 3.9 Với mức ý nghĩa $\alpha = 0{,}05$; kết luận chấp nhận giả thuyết H_o. Kết quả kiểm định cho thấy không có hiệu ứng ARCH trên chuỗi nhiễu bình phương. Kết quả này phù hợp với nhận định trên chuỗi số liệu và đồ thị tự tương quan của chuỗi bình phương nhiễu.

Kiểm định hiệu ứng ARCH trên các chuỗi số liệu SAM đều cho thấy không xuất hiện hiệu ứng ARCH. Nhóm tác giả của đề tài thực hiện kiểm định hiệu ứng ARCH trên chuỗi số liệu $\left\{r_{5t};\, t = \overline{1,\left[\frac{n}{5}\right]}\right\}$ thay vì chuỗi $\{r_t;\, t = \overline{1,n}\}$ (số 5 chỉ số ngày giao dịch trong tuần).

<u>**Mô hình 3.**</u> Xét chuỗi $\left\{y_t = r_{5t};\, t = \overline{1,\left[\frac{n}{5}\right]}\right\}$ với n = 983, chuỗi $\{r_t;\, t = \overline{1,n}\}$ lãi suất của chứng khoán SAM cho ở ví dụ 2.

Số số liệu: 196

Nguồn gốc số liệu: *http://www.hse.org.vn*

Ước lượng tham số bằng mô hình:

+) Ước lượng tham số bằng mô hình AR(0) cho chuỗi $\{y_t\}_t$: $y_t = \varepsilon_t$

+) Dùng thống kê Ljung – Box kiểm định hiệu ứng ARCH trên chuỗi $\{\varepsilon_t\}_t$

h	Q_{LB}	$\chi^2(h), \alpha = 0,05$
1	1,576	3,841
2	7,228	5,991
3	10,426	7,815
4	14,72	9,488
5	18,335	11,071
6	20,958	12,592
7	21,655	14,067
8	21,872	15,507
9	24,29	16,919
10	24,448	18,307

Bảng 3.15. Giá trị thống kê Q_{LB} và giá trị tới hạn $\chi^2(h)$

Với mức ý nghĩa $\alpha = 0,05$; đều bác bỏ giả thuyết H_o và chấp nhận giả thuyết H_1 có hiệu ứng ARCH trên chuỗi số liệu. Kết quả số này cũng phù hợp với sự tạo cụm biến động trên đồ thị chuỗi lãi suất và chuỗi bình phương lãi suất, hiện tượng tự tương quan trên đồ thị hàm tự tương quan và tự tương quan riêng của chuỗi lãi suất và chuỗi bình phương lãi suất.

+) Nhận dạng mô hình ARCH(4).

$$\varepsilon_t = \sqrt{h_t}.\upsilon_t;\ \upsilon_t \sim \text{IID}(0;\ 1)$$

$$\begin{aligned} h_t &= 0,00272 + 0,047122\,\varepsilon_{t-1}^2 + 0,138627\varepsilon_{t-2}^2 \\ &\quad + 0,097368\varepsilon_{t-3}^2 + 0,109213\varepsilon_{t-4}^2 \end{aligned} \qquad (3.22)$$

+) Kiểm định sự phù hợp của mô hình. Thống kê Q_{LB} cho chuỗi chuẩn hoá $\{\upsilon_t\}$.

h	Q_{LB}	$\chi^2(h), \alpha = 0{,}05$
1	1,225	3,841
2	2,866	5,991
3	4,197	7,815
4	4,623	9,488
5	5,63	11,071
6	5,631	12,592
7	5,641	14,067
8	5,665	15,507
9	5,885	16,919
10	5,897	18,307

Bảng 3.16. Giá trị thống kê Q_{LB} và giá trị tới hạn $\chi^2(h)$

So sánh giá trị thống kê Q_{LB} với giá trị tới hạn khi bình phương ở bảng 3.11. Với mức ý nghĩa $\alpha = 0{,}05$; kết luận mô hình (3.47) là phù hợp.

Mô hình 4. Xét chuỗi $\mathcal{Y}_1 = \left\{ y_t = r_{5(t-1)+1};\ t = \overline{1, \left[\frac{n}{5}\right]} \right\}$;

$\mathcal{Y}_2 = \left\{ y_t = r_{5(t-1)+2};\ t = \overline{1, \left[\frac{n}{5}\right]} \right\}$; $\mathcal{Y}_3 = \left\{ y_t = r_{5(t-1)+3};\ t = \overline{1, \left[\frac{n}{5}\right]} \right\}$

và $\mathcal{Y}_4 = \left\{ y_t = r_{5(t-1)+4};\ t = \overline{1, \left[\frac{n}{5}\right]} \right\}$ với n = 983, chuỗi $\{r_t;\ t = \overline{1,n}\}$ lãi suất của chứng khoán SAM cho ở ví dụ 2.

Số số liệu: 196

Nguồn gốc số liệu: *http://www.hse.org.vn*

Ước lượng tham số bằng mô hình:

+) Ước lượng tham số bằng mô hình AR(0) cho các chuỗi $\mathcal{Y}_t$ $\left(t = \overline{1,4}\right)$: $y_t = \varepsilon_t$

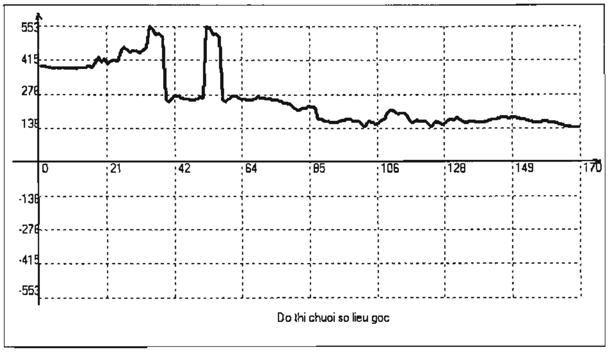
+) Với mức ý nghĩa $\alpha = 0{,}05$; dùng thống kê Ljung – Box kiểm định hiệu ứng ARCH trên các chuỗi $\{\varepsilon_t\}$ thu được đều chấp nhận giả thuyết H_o, không có hiệu ứng ARCH trên chuỗi số liệu.

➢ ***Kết luận 2.*** Các kết quả thu được ở ví dụ 4 thật bất ngờ là không xuất hiện hiệu ứng ARCH trên dãy các số liệu. Các số liệu theo tuần ở mô hình 4 đều là các số liệu lấy vào ngày thứ 2, thứ 3, thứ 4 và thứ 5 hàng tuần. Trong khi đó, đối với ví dụ 3, số liệu lấy vào ngày thứ 6 hàng tuần (ngày giao dịch cuối tuần). Như vậy đối với chuỗi SAM, xuất hiện hiệu ứng ARCH đối với chuỗi lãi suất theo tuần (số liệu lấy theo ngày thứ 6 hàng tuần).

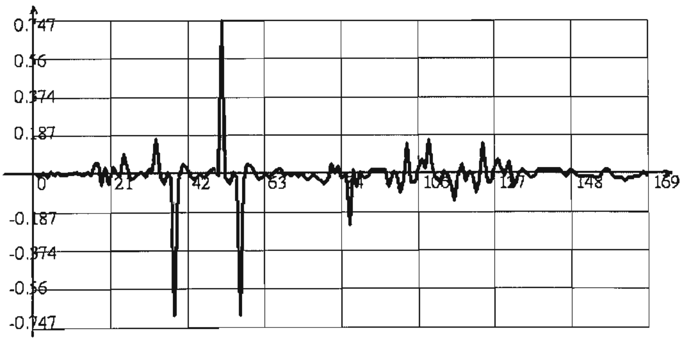
Mô hình 5. Chuỗi DHG (chỉ số chứng khoán của công ty Cổ phần dược Hậu Giang).

Số số liệu: 170 (Thời điểm: 9/2007 đến 7/2008)

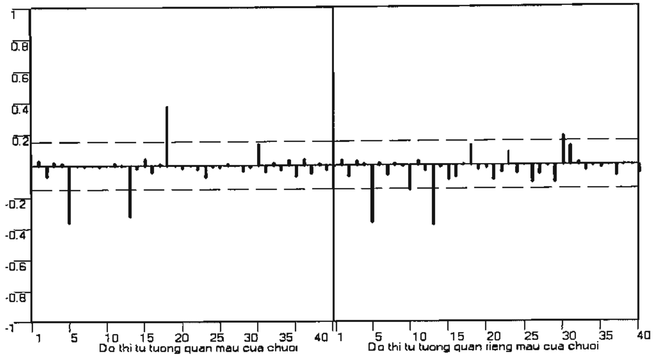
Nguồn số liệu:*http://www.ssi.com.vn*



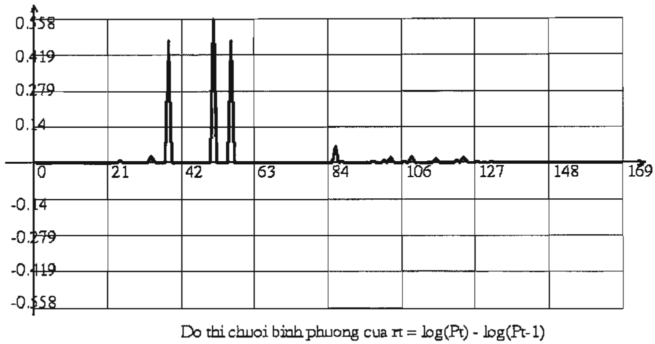
Hình 3.17. Đồ thị chuỗi chỉ số chứng khoán DHG



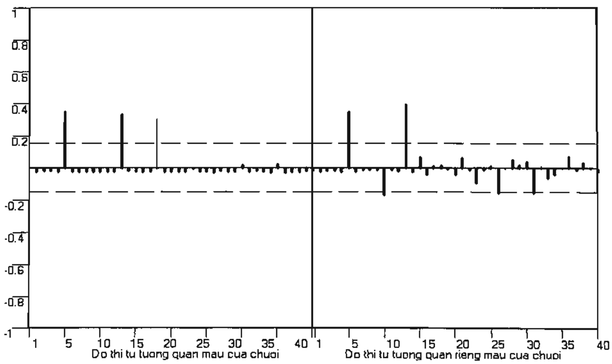
Hình 3.18. Đồ thị chuỗi lãi suất chứng khoán DHG



Hình 3.19. Đồ thị tự tương quan và tự tương quan riêng của chuỗi lãi suất.



Hình 3.20. Đồ thị chuỗi bình phương lãi suất chứng khoán.



Hình 3.21. Đồ thị tự tương quan và tự tương quan riêng chuỗi bình phương lãi suất chứng khoán.

Hiện tượng tạo cụm biến động ít và tính ít tương quan của chuỗi bình phương lãi suất chứng khoán cho thấy có thể không có hiệu ứng ARCH đối với chuỗi lãi suất. Ước lượng tham số bằng mô hình:

+) Ước lượng tham số bằng mô hình ARMA(1, 1) cho chuỗi $\{r_t\}_t$

$$r_t - 0,6092678\, r_{t-1} = -\,0,00232 + \varepsilon_t - 0,578998\,\varepsilon_{t-1}$$

+) Dùng thống kê Ljung – Box kiểm định hiệu ứng ARCH trên chuỗi $\{\varepsilon_t\}_t$

h	Q_{LB}	$\chi^2(h), \alpha = 0,05$
1	16,347	3,841
2	16,603	5,991
3	16,609	7,815
4	17,344	9,488
5	34,946	11,071
6	36,404	12,592
7	36,45	14,067
8	36,741	15,507
9	37,116	16,919
10	37,448	18,307

Bảng 3.22. Giá trị thống kê Q_{LB} và giá trị tới hạn $\chi^2(h)$

So sánh giá trị thống kê Q_{LB} với giá trị tới hạn khi bình phương ở bảng 3.9 Với mức ý nghĩa $\alpha = 0,05$; kết luận bác bỏ giả thuyết H_0. Kết quả kiểm định cho thấy có hiệu ứng ARCH trên chuỗi nhiễu bình phương. Kết quả này phù hợp với nhận định trên chuỗi số liệu và đồ thị tự tương quan của chuỗi bình phương nhiễu.

+) Nhận dạng mô hình ARCH(1).

$$\varepsilon_t = \sqrt{h_t}.\upsilon_t;\ \upsilon_t \sim \mathrm{IID}(0;\ 1)$$

$$h_t = 0,010665 + 0,31194\ \varepsilon_{t-1}^2 + 0,138627\,\varepsilon_{t-2}^2 \qquad (3.23)$$

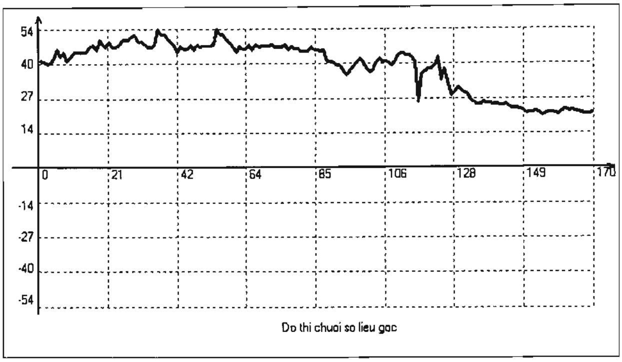
+) Kiểm định sự phù hợp của mô hình. Thống kê Q_{LB} cho chuỗi chuẩn hoá $\{\upsilon_t\}$ cho kết quả.

➢ ***Kết luận 3.*** Các kết quả kiểm định đối với mô hình 5 cho thấy có hiệu ứng ARCH trên chuỗi DHG. Kết quả này phù hợp với nhận định trên chuỗi số liệu và đồ thị tự tương quan của chuỗi bình phương nhiễu. Từ đó ta nhận dạng được mô hình cho biến động h_t là mô hình (3.23).

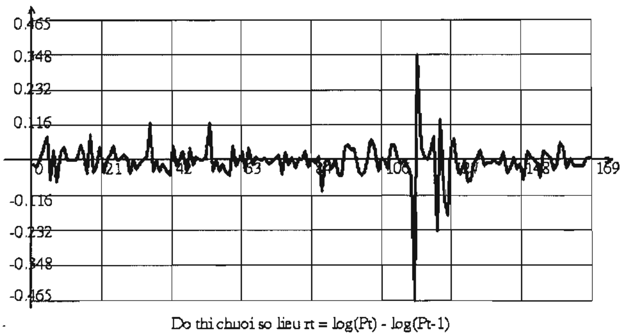
Mô hình 6. Chuỗi DHC (chỉ số chứng khoán của công ty Cổ phần thuỷ điện Ry Ninh II).

Số số liệu: 170 (Thời điểm: 9/2007 đến 7/2008)

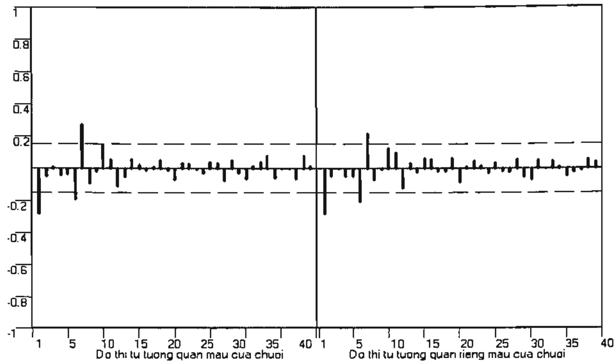
Nguồn số liệu:*http://www.ssi.com.vn*



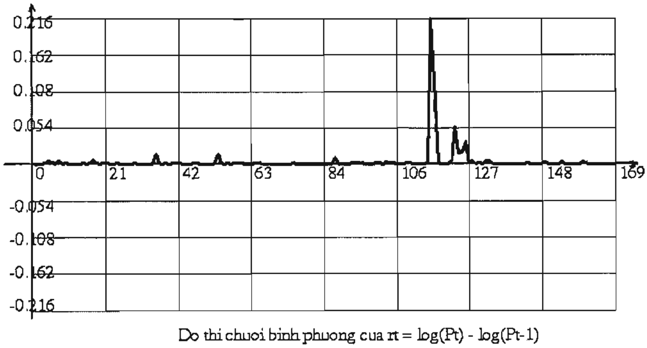
Hình 3.23. Đồ thị chuỗi chỉ số chứng khoán DHC



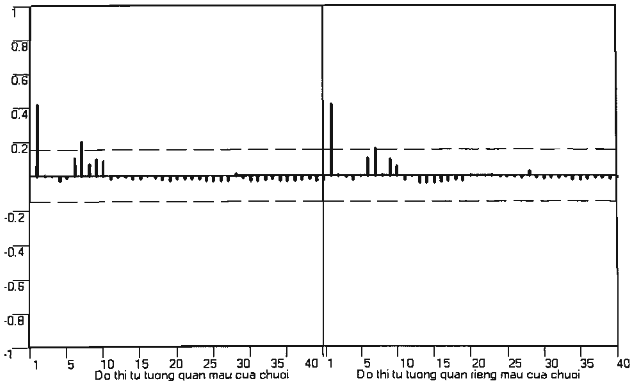
Hình 3.24. Đồ thị chuỗi lãi suất chứng khoán DHG



Hình 3.25. Đồ thị tự tương quan và tự tương quan riêng của chuỗi lãi suất.



Hình 3.26. Đồ thị chuỗi bình phương lãi suất chứng khoán.



Hình 3.27. Đồ thị tự tương quan và tự tương quan riêng chuỗi bình phương lãi suất chứng khoán.

Hiện tượng tạo cụm biến động ít và tính ít tương quan của chuỗi bình phương lãi suất chứng khoán cho thấy có thể không có hiệu ứng ARCH đối với chuỗi lãi suất. Ước lượng tham số bằng mô hình:

+) Ước lượng tham số bằng mô hình ARMA(1, 1) cho chuỗi $\{r_t\}_t$

$$r_t - 0{,}6092678\, r_{t-1} = -\,0{,}00232 + \varepsilon_t - 0{,}578998\,\varepsilon_{t-1}$$

+) Dùng thống kê Ljung – Box kiểm định hiệu ứng ARCH trên chuỗi $\{\varepsilon_t\}_t$

h	Q_{LB}	$\chi^2(h), \alpha = 0{,}05$
1	0,606	3,841
2	0,684	5,991
3	0,75	7,815
4	0,757	9,488
5	0,759	11,071
6	1,263	12,592
7	3,796	14,067
8	3,797	15,507
9	3,965	16,919
10	10,651	18,307

Bảng 3.28. Giá trị thống kê Q_{LB} và giá trị tới hạn $\chi^2(h)$

➢ ***Kết luận 4.*** So sánh giá trị thống kê Q_{LB} với giá trị tới hạn khi bình phương ở bảng 3.28. Với mức ý nghĩa $\alpha = 0{,}05$; kết luận chưa có cơ sở để bác bỏ giả thuyết H_o. Kết quả kiểm định cho thấy không có hiệu ứng ARCH trên chuỗi nhiễu bình phương. Kết quả này phù hợp với nhận định trên chuỗi số liệu và đồ thị tự tương quan của chuỗi bình phương nhiễu. Do đó đối với chuỗi thời gian DHC không tìm thấy hiệu ứng ARCH.

3.2.4. Dự báo biến động của các chuỗi thời gian tài chính theo mô hình ARCH(Q)

Sử dụng các kết quả thu được của nhận dạng mô hình ARCH(Q) đối với chuỗi Nasdaq và chuỗi SAM, DHG. Đối với chuỗi Nasdaq, thực hiện dự báo 12 giá trị của biến động của chuỗi lãi suất theo ngày. Với chuỗi SAM, thực hiện dự báo 12 giá trị của biến động lãi suất theo tuần. Với chuỗi DHG, thực hiện dự báo 12 giá trị của biến động theo ngày.

Hàm dự báo cho các mô hình

* **Với chuỗi Nasdaq:** Với mô hình cho ở công thức (3.21)

$$\varepsilon_t = \sqrt{h_t}.\upsilon_t;\ \upsilon_t \sim IID(0;\ 1)$$

$$h_t = 0{,}000048 + 0{,}13719\varepsilon_{t-1}^2 + 0{,}199087\varepsilon_{t-2}^2$$

ta xây dựng được hàm dự báo như sau:

$$f_{n,k} = 0{,}000048 + 0{,}13719.f_{n,k-1} + 0{,}199087.f_{n,n-2} \quad (3.24)$$

Ở đây, $f_{n,k-i} = \varepsilon_{n+k-i}^2$ nếu $k - i \le 0$.

* **Với chuỗi SAM:** Với công thức cho ở (3.22)

$$\varepsilon_t = \sqrt{h_t}.\upsilon_t;\ \upsilon_t \sim IID(0;\ 1)$$

$$h_t = 0{,}00272 + 0{,}047122\ \varepsilon_{t-1}^2 + 0{,}138627\varepsilon_{t-2}^2 + 0{,}097368\varepsilon_{t-3}^2 + 0{,}109213\varepsilon_{t-4}^2$$

ta xây dựng được hàm dự báo như sau:

$$f_{n,k} = 0{,}00272 + 0{,}047122.f_{n,k-1} + 0{,}138627.f_{n,k-2} + 0{,}097368.f_{n,k-3} + 0{,}109213.f_{n,k-4} \quad (3.25)$$

Ở đây, $f_{n,k-i} = \varepsilon_{n+k-i}^2$ nếu $k - i \le 0$.

* **Với chuỗi DHG:** Với công thức (3.23)

$$\varepsilon_t = \sqrt{h_t}.\upsilon_t;\ \upsilon_t \sim IID(0;\ 1)$$

$$h_t = 0{,}010665 + 0{,}31194\ \varepsilon_{t-1}^2 + 0{,}138627\,\varepsilon_{t-2}^2$$

ta xây dựng được hàm dự báo như sau:

$$f_{n,k} = 0{,}010665 + 0{,}31194.f_{n,k-1} + 0{,}138627.f_{n,k-2} \quad (3.26)$$

Ở đây, $f_{n,k-i} = \varepsilon_{n+k-i}^2$ nếu $k - i \le 0$.

Áp dụng các công thức (3.24), (3.25) và (3.26) ta tính được các giá trị dự báo cho biến động của các chuỗi giá chứng khoán Nasdaq, SAM, DHG.

Dự báo của chuỗi thời gian Nasdaq, SAM, DHG

	Kết quả 1	*Kết quả 2*	*Kết quả 3*
Dự báo h bước	Mô hình (3.21)	Mô hình (3.22)	Mô hình (3.23)
1	0,000049	0,002833	0,011285
2	0,000055	0,002879	0,014421
3	0,000066	0,003261	0,015292
4	0,000068	0,003562	0,015535
5	0,000071	0,00393	0,015602
6	0,000072	0,004031	0,015602
7	0,000072	0,004158	0,015621
8	0,000073	0,004247	0,015626
9	0,000073	0,004318	0,015627
10	0,000073	0,004358	0,015628
11	0,000073	0,004392	0,015628
12	0,000073	0,004416	0,015628

Bảng 3.29. Kết quả dự báo biến động của chỉ số chứng khoán theo các mô hình (3.21), (3.22) và (3.23)

Chú ý: Đối với chuỗi số liệu SAM ta chỉ thực hiện dự báo biến động giá chứng khoán vào ngày thứ 2 hàng tuần. Cũng phương pháp như vậy ta áp dụng cho các ngày thứ 3, 4, 5, 6 hàng tuần.

Kết luận 5: Như vậy, với việc phát hiện ra bằng chứng ARCH (hiệu ứng của chuỗi thời gian tài chính) dù phát hiện dưới bất kỳ hình thức nào về số liệu ta đều sử dụng được các kết quả của việc nhận dạng của mô hình đó vào bài toán dự báo biến động.

Kết luận 6: Thực hiện đánh giá phương pháp dự báo theo mô hình ARCH đối với 2 chuỗi số liệu trên với phương pháp NF ta sử dụng biến động thực tế $v_t = (r_t - \bar{r})^2$, $t = \overline{1,n}$; $\bar{r}$ là trung bình của chuỗi lãi suất. Hoặc thực hiện đánh giá sai số của dự báo dựa trên phương pháp đánh giá sai số ngoài mẫu đều cho thấy các kết quả dự báo trên đều đáng tin cậy.

3.3. Kết luận và kiến nghị - đề xuất giải pháp

3.3.1. Kết luận

Với cách tiếp cận tương tự như đối với các chuỗi Nasdaq, SAM, DHG có thể thực hiện kiểm định bằng chứng của hiệu ứng ARCH/ GARCH trên các chuỗi số liệu chứng khoán Việt Nam, từ đó áp dụng vào bài toán dự báo biến động chứng khoán theo mô hình ARCH/ GARCH. Quá trình phát hiện hiệu ứng ARCH/ GARCH trên số liệu chứng khoán Việt Nam gặp nhiều khó khăn, phức tạp so với việc phát hiện hiệu ứng đó trên thị trường chứng khoán của thị trường tài chính thế giới (Mỹ, Anh, Nhật, ...). Quá trình nhận dạng mô hình ARCH/ GARCH và ứng dụng mô hình vào dự báo biến động một lần nữa khẳng định ảnh hưởng của phương sai có điều kiện tới quá trình phát sinh và lan truyền các biến động của chuỗi thời gian tài chính mà các quá trình ARMA không thể giải thích được. Mô hình ARCH/ GARCH đã giải thích được nguồn gốc của hiện tượng tạo các cụm biến động trong các chuỗi thời gian tài chính, trong khi đó thực hiện các kiểm định hiệu ứng ARCH/ GARCH trên các chuỗi số liệu A, B, C, D, E, F, G [15] của Box – Jenskin thuộc các lĩnh vực khoa học tự nhiên, kỹ thuật thì không thấy xuất hiện hiệu ứng ARCH/ GARCH. Đây chính là thế mạnh của mô hình ARCH/ GARCH trong lĩnh vực nghiên cứu chuỗi thời gian tài chính.

Thị trường chứng khoán Việt Nam tuy mới ra đời (năm 2000) và rất còn non trẻ so với các thị trường khác, nhưng với hơn 1000 phiên giao dịch bắt đầu từ phiên giao dịch đầu tiên 28 tháng 7 năm 2000 cũng đã cung cấp cho các nhà nghiên cứu một khối lượng lớn số liệu đủ để phân tích, ước lượng và dự báo. Để thấy được tầm quan trọng của các mô hình ARCH và GARCH trong việc giải thích được các đặc trưng của biến động trên các chuỗi số liệu chứng khoán Việt Nam, đòi hỏi nhiều thời gian và công sức của các nhà nghiên cứu về mặt lý thuyết cũng như mặt ứng dụng. đề tài chỉ bước đầu tiếp cận với mô hình ARCH và tìm tòi cách áp dụng trên các chuỗi thời gian tài chính cụ thể nhằm phát hiện hiệu ứng ARCH, ước lượng các tham số của mô hình, kiểm định sự phù hợp của mô hình và ứng dụng mô hình đó vào bài toán dự báo.

Tóm lại, đề tài này đạt được những kết quả sau đây:

- *Một là,* đề tài đã tổng kết các phương pháp dự báo của chuỗi thời gian, phân tích các khía cạnh của công tác dự báo. Từ đó, tạo nên cơ sở lý luận vững chắc để áp dụng phương pháp dự báo định lượng vào giải quyết các bài toán thực tiễn của kinh tế đạ ra nói chung và lĩnh vực tài chính nói riêng.

Hai là, đề tài áp dụng kỳ vọng có điều kiện chính là phép chiếu trong không gian Hilbert, xây dựng được công thức dự báo điểm cho các mô hình chuỗi thời gian ARMA, ARARMA hoặc ARAR và ARCH. Đặc biệt, xây dựng công thức dự báo điểm cho quá trình ARMA khi có hiệu ứng ARCH. Từ đó, xây dựng được các thuật toán dự báo theo các mô hình tương ứng. Từ đây, chúng ta có một cơ sở lý thuyết vững chắc cho việc xây dựng các hàm dự báo mà hầu hết các bài toán dự báo định lượng đều phải sử dụng các hàm dự báo này.

Ba là, đề tài đã ứng dụng thành công các kết quả lý thuyết vào các bài toán cụ thể: Bài toán dự báo biến động giá của chứng khoán và tiếp cận bài toán dự báo biến động giá chứng khoán theo mô hình ARCH, sử dụng kết quả của bài toán nhận dạng mô hình ARCH bằng phương pháp bình phương cực tiểu theo thuật toán 2.5 sau khi kiểm định hiệu ứng ARCH. Việc tìm kiếm hiệu ứng ARCH/ GARCH trên thị trường chứng khoán Việt Nam đòi hỏi các tác giả của đề tài phải bỏ nhiều công sức lựa chọn và phân loại các tập số liệu thành nhiều loại hoặc có thể phân thành nhiều phần để kiểm tra hiệu ứng ARCH/ GARCH. Đây là một điều kiện kiên quyết, một chuỗi giá chứng khoán có hiệu ứng ARCH/ GARCH thì mới có thể dùng công cụ của chuỗi thời gian tài chính để dự báo biến động giá của chứng khoán đó. Sau khi nhận dạng và áp dụng các phương pháp dự báo chuỗi thời gian, đề tài cung cấp 3 kết quả số cụ thể về các giá trị dự báo biến động của các chuỗi giá chứng khoán NASDAQ, SAM, DHG. Các kết luận về kết quả số đó là các kết luận 1 đến kết luận 6 của mục 3.2.

Bốn là, cùng với việc xây dựng các kết qủa về lý thuyết dự báo bằng mô hình chuỗi thời gian đề tài còn xây dựng phần mềm phân tích chuỗi thời gian bằng ngôn ngữ Visual Basic, nhằm giúp thực hiện các thao tác và thuật toán trong đề tài. Phần mềm này tạo điều kiện thuận lợi cho việc phân tích chuỗi thời gian, đặc biệt là chuỗi thời gian tài chính, ... Để chứng minh tính đúng đắn của phân mềm, các tác giả của đề tài sử dụng 2 phương pháp kiểm tra: phương pháp đối sánh với các phầm mềm chuyên dụng như Eviews, Statas, SPS, ... và phương pháp mô phỏng. Các kết quả kiểm tra đều chứng minh tính đúng đắn của các kết quả số do phần mềm cung cấp.

3.3.2. Kiến nghị - đề xuất giải pháp

Để thực sự thấy được tầm quan trọng của các công cụ phân tích định lượng trong các lĩnh vực tài chính cũng như ứng dụng của đề tài trong công tác giảng dạy và NCKH tại trường Đại học Ngoại thương, nhóm tác giả đề tài xin đề xuất một số giải pháp mang tính gợi mở sau đây:

Thứ nhất, cần có sự phối hợp giữa các công cụ định lượng với các công cụ chuyên môn của từng chuyên ngành trong Nhà trường. Hay nói cách khác là sự bắt tay trao đổi để tìm tiếng nói chung trong giảng dạy và NCKH. Chẳng hạn như giữa các môn Toán với các môn Tài chính tiền tệ, Kinh tế vi mô, vĩ mô, Lý tuyết trò chơi, ...

Thứ hai, cần đổi mới mạnh mẽ việc giảng dạy các môn Toán, Toán ứng dụng trong Nhà trường để đảm bảo tính thiết thực và tính ứng dụng. Mạnh dạn bỏ những nội dung mang tính lý thuyết kinh viện và thay vào đó là đưa vào các nội dung ứng dụng thiết thực vào các lĩnh vực kinh tế, tài chính như chuỗi thời gian, phương trình sai phân tích phân, quá trình ngẫu nhiên, lý thuyết trò chơi, ...

Thứ ba, để thực hiện tốt chức năng ứng dụng của đề tài, cũng như thực hiện nhiệm vụ là đổi mới nội dung, chương trình, phương pháp giảng dạy các môn học ứng dụng, nên chăng đối với các môn học mà nội dung của nó gắn với các phần mềm chuyên dụng, các tổ chuyên môn nên quy định một thời lượng tối thiểu cho việc giới thiệu ứng dụng của các phần mềm này vào nội dung của môn học.

Thứ tư, chất lượng của các mô hình dù nhận dạng ở bất kỳ phần mềm nào cũng phụ thuộc vào nguồn gốc số liệu, đây là một bài toán ở tầm vĩ mô, nhưng trong điều kiện giảng dạy và NCKH, cũng nên cần xây dựng một thư viện số liệu thuộc các lĩnh vực nhằm phục vụ tốt cho công tác giảng dạy và NCKH ở trong một trường ĐH khối kinh tế hoặc cần xây dựng mối quan hệ hợp tác với các trung tâm phân tích số liệu có uy tín trong nước và quốc tế.

3.4. Những vấn đề còn tồn tại

Tiếp nối những vấn đề trong đề tài, các vấn đề mở sau cần được tiếp tục nghiên cứu và giải quyết:

-*Vấn đề 1*, sử dụng phương pháp bình phương cực tiểu ước lượng tham số của mô hình GARCH, ARCH – M , E – GARCH, CHARMA [36].

- *Vấn đề 2*, nhận dạng các mô hình GARCH, ARCH – M, E – GARCH cho tất cả chuỗi số liệu chứng khoán Việt Nam, từ năm 2000 – 2006. Từ đó, áp dụng vào bài toán dự báo biến động.

- *Vấn đề 3*, các mô hình biến động ngẫu nhiên (the Stochastic volatility model), mô hình biến động ngẫu nhiên bộ nhớ dài (The long – memory stochastic volatility model) và áp dụng vào thị trường chứng khoán Việt Nam.

- *Vấn đề 4*, Các mô hình chuỗi thời gian phi tuyến và ứng dụng vào phân tích thị trường chứng khoán Việt Nam.

Để tăng tính hiệu quả của đề tài hơn nữa cũng như các bài toán ứng dụng, nhóm tác giả của đề tài kỳ vọng rằng trong các vấn đề đặt ra ở đây, nếu có sự hợp tác chặt chẽ của những người làm toán với các nhà kinh tế thì tính thiết thực của các bài toán đặt ra ở đây tăng lên gấp bội.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Nguyễn Văn Hữu, Nguyễn Hữu Dư (2003), *Phân tích thống kê và dự báo,* NXB ĐHQG Hà nội, Hà nội.
2. Tống Đình Quỳ (2003), *Giáo trình xác suất thống kê*, NXB ĐHQG Hà nội, Hà nội.
3. Nguyễn Hồ Quỳnh (2004), *Chuỗi thời gian - phân tích và nhận dạng*, NXB Khoa học và kỹ thuật, Hà nội.
4. Nguyễn Duy Tiến, Đặng Hùng Thắng (2001), *Các mô hình xác suất và ứng dụng (Phần II)*, NXB ĐHQG Hà nội, Hà nội.
5. Nguyễn Duy Tiến, Đặng Hùng Thắng (2001*), Các mô hình xác suất và ứng dụng (Phần III)*, NXB ĐHQG Hà nội, Hà nội.
6. Nguyễn Duy Tiến, Vũ Việt Yên (2000), *Lý thuyết xác suất*, Nhà xuất bản Giáo dục.
7. Hoàng Tuỵ (2005), *Hàm thực và giải tích hàm*, NXB ĐHQG Hà nội, Hà nội.
8. Nguyễn Lê Văn (2005), *Xử lý mùa trong mô hình ARIMA*, Đồ án tốt nghiệp đại học, Trường Đại học Bách Khoa Hà nội, Hà nội.
9. Võ Văn Vinh (2005), *Mô hình ARCH và chuỗi thời gian tài chính*, Đồ án tốt nghiệp đại học, Trường Đại học Bách Khoa Hà nội, Hà nội.
10. Vương Quân Hoàng (2004), "Hiệu ứng GARCH trên dãy lợi suất thị trường chứng khoán Việt Nam 2000 – 2003 ", *Tạp chí ứng dụng Toán học, tập II, số 1,* trang 15 -30.

Tiếng Anh

11. Bell W. R. (1984),"A Introduction to Forecasting with Time Series Models ", *Mathematics and Economics*, 3, pp. 241 – 255, North – Holland.
12. Bollerslev T. (1986), "Generalised autoregressive conditonal heteroskedasticity ", *Journal of econmetrics*, 31, pp. 307 – 327.
13. Bera A. K., Higgins M . L. (1993), "ARCH models: Properties, Estimation and Testing ", *Journal of Economic Surveys*, 7 (4), pp. 305 –362.

14. Bose A., Mukherjee K. (2003), “Estimating the ARCH parameters by solving linear equations ”, *Journal of Time Series Analysis*, 24(2), pp. 127 - 137.
15. Box G. E. P. , Jenkins G.M (1970), *Time Series Analysis forecasting and control*, Holden – Day.
16. Box G. E. P., Pierce (1970), “Distribution of Autocerralation in auoregressive – integrated Moving Averge Time Series Models ”, *Journal of the American Statistical Association*, 65(332), pp. 1509 – 1562.
17. Brockwell P. J., Davis R. A. (1996), *Introduction to Time Series and Forecasting*, Springer, NewYork.
18. Chong C. W., Ahmad M. I. and Abdullah M. Y.(1999), “Performance of GARCH Models in Forecasting Stock Market Volatility ”, *Journal of forecating*, 18, pp. 333 – 343.
19. Engle R. F. (1982), “Autoregressive Conditional Heteroskedasticity with Estimates of Variance of U.K Inflation ”, *Econometrica*, 50, pp. 987 – 1008.
20. Hall P. , Heyde C. (1980), *Martingale Limit Theory and its Application*, Academic Press, New York.
21. Hamilton J. D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.
22. Jorion P. (1995), “Predicting volatility in the foreign exchange market”, *Journal of Finance, 50*, pp. 507 – 528.
23. Li W. K., Ling S., McAleer M. (2003), “Recent Theoretical Results for Time Series Models with GARCH errors”, *Journal of Economic Surveys*, 16(3), pp. 245- 255.
24. Ljung G. M. , Box G. E. P (1978), “On a measure of lack of fit in tiem series models”, *Biometrika, 65*, pp. 297 -303.
25. Makridakis S., Wheelwright S. C. , Hydman R. J. (1998), *Forecasting – Methods and Application*, John Wiley & Sons, Inc, NewYork.
26. Mann H. B., (1943), "On Stochastic Limit and Order Relationships", *The Annals of Mathematical Statistics, 14*, pp.217 – 226.

27. McLeod A. I. (1978), "On the Distribution of Residual Autocorrelations in Box – Jenkins Models", *Journal of the Royal Statistical Society*. Series B (Methodological), 40 (3), pp. 296 – 302.

28. McLeod A. I. and Li W.K. (1983), "Diagnostic checking ARMA Time Series Models using Squared – Residual Autocorrelations", *Journal of Time Series Analysis*, 4, pp. 269 – 273.

29. Monti A. C. (1994), "A Proposal for a Residual Autocorrelation Test in Linear Models ", *Biometrika*, 81(4), pp. 776 – 780.

30. Najand M. (2002), "Forecasting Stock Index Futures Price Volatility: Linear vs. Nonlinear Models ", *The financial Review*, 37, pp. 93 -104.

31. Parzen E. (1982), "ARARMA Models for Time Series Analysis and Forecasting", *Journal of Forecasting*, 1, pp. 67 – 82.

32. Parzen E. (2002),"Time Series Analysis, Sieve, Memory, and Exponential Spectral Models ", *Dedicated to the 40th Anniversary of the Department of Statistics at Texas A & M University*, Department of Statistics, Texas A & M University.

33. Pollock D. S. G. (1999), *A handbook of Time Series Analysis*, Signal Processing and Dynamics, Acedemic Press, New York.

34. Poon S. , Granger C. J. W. (2003), "Forecasting volatility in financial markets: A review ", *Journal of Economic Liturature, 41*, pp.478 – 639.

35. Shorack G. R. (2000), *Probability for Statisticians*, Springer -Verlag, NewYork.

36. Tsay R.S. (2002), *Analysis of Financial Time Series*, A Wiley – Interscience publication, John Wiley & Sons, Inc, New York.

37. Stoll H., Whaley R. E. (1990), " Stock Market Structure and Volatility ", *The Review of Financial Studies*, 3(1), pp. 37 -71.

38. www – personal.buseco.monash.edu.au.

39. www.vcbs.com.vn.