

Ứng dụng mô hình kinh tế lượng trong dự báo giá chứng khoán

DƯƠNG NGÂN HÀ

Học viện Ngân hàng

Thị trường chứng khoán được coi là một kênh đầu tư hấp dẫn có thể mang lại tỷ suất sinh lời cao tuy nhiên lại chứa đựng rất nhiều rủi ro. Giá thị trường của cổ phiếu biến động không ngừng, phụ thuộc chủ yếu vào cung cầu của chính cổ phiếu đó và xu thế chung của thị trường. Tìm kiếm giá trị thực và xu thế biến động giá của cổ phiếu luôn là mục đích mà các nhà đầu tư hướng tới. Do vậy, việc đưa ra dự báo về xu hướng biến động của giá chứng khoán được coi là một chiến lược phù hợp cho hoạt động đầu tư chứng khoán trên thị trường. Một trong số các phương pháp dự báo giá chứng khoán là dựa trên việc áp dụng mô hình ARIMA trong kinh tế lượng để đưa ra dự báo cho chuỗi giá chứng khoán. Bài viết đề cập tới cách thức sử dụng mô hình ARIMA trong việc dự báo giá chứng khoán cụ thể là chỉ số VN-Index, trên cơ sở đó có thể mở rộng việc dự báo cho các chứng khoán đang giao dịch trên thị trường chứng khoán Việt Nam.

Giới thiệu



Được biết đến chính thức khi Trung tâm Giao dịch Chứng khoán Tp.Hồ Chí Minh giao dịch phiên đầu tiên vào ngày 22/7/2000, thị trường chứng khoán đã và đang trở thành một kênh đầu tư hấp dẫn đối với các nhà đầu tư từ tổ chức tới cá nhân. Sau 12 năm phát triển, thị trường chứng

khoán đã trải qua rất nhiều thăng trầm thể hiện rõ qua sự biến động của chỉ số VN-Index. Có thời điểm VN-Index đã đạt tới 1.170 điểm vào ngày 12/3/2007 và sau đó chạm đáy ở mức 235 điểm tại ngày 24/2/2009. Sự biến động mạnh của chỉ số mang lại mức sinh lời cao cho các chủ thể đầu tư nhưng cũng tồn tại rất nhiều rủi ro bởi không



phải lúc nào nhà đầu tư cũng dự báo chính xác xu hướng của giá cổ phiếu trong tương lai. Hơn thế nữa, đo lường sự biến động của chỉ số VN-Index cũng đồng thời đo lường rủi ro hệ thống của thị trường, vì vậy việc dự báo được sự tăng giảm của VN-Index cũng sẽ giúp cho các nhà đầu tư nhận biết chiều hướng biến động giá của các cổ phiếu trên thị trường.

Phương pháp

Mô hình ARIMA (Autoregressive Integrated Moving Average- trung bình trượt tự hồi quy tích hợp) hay còn được biết đến với tên gọi phương pháp Box-Jenkins (BJ) không những nhấn mạnh vào việc xây dựng một phương trình hay hệ phương trình đồng thời mà còn chú ý vào việc phân tích tính ngẫu nhiên, tính chất chuỗi số liệu trên cơ sở chuỗi dữ liệu quá khứ và đưa ra dự báo cho sự biến động của chuỗi thời gian. Phương pháp Box-Jenkins bao gồm 4 bước: Định dạng mô hình, ước lượng, kiểm định tính thích hợp của mô hình và dự báo.

Các phương pháp này liên quan tới biến số thay đổi theo thời gian và phụ thuộc vào thời gian hiện tại và giá trị quá khứ (có nghĩa là không phụ thuộc vào bất kỳ biến số nào bên ngoài). Nếu Y_t là giá trị của biến số tại thời điểm t , ta có phương trình:

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots, Y_0, t)$$

Có nghĩa là giá trị của biến số tại thời điểm t đơn thuần là phương trình giá trị trước đó và thời gian, không có biến số

nào liên quan. Mục đích của phương pháp này là tìm ra bản chất của phương trình f trên cơ sở dự báo giá trị cho Y_t .

Theo Box- Jenkin mọi quá trình ngẫu nhiên có tính dừng đều có thể biểu diễn bằng mô hình Trung bình trượt tự hồi quy tích hợp ARIMA.

- Mô Hình Tự Hồi Quy Bậc p - AR(p): Trong mô hình tự hồi quy quá trình phụ thuộc vào tổng có trọng số của các giá trị quá khứ và số hạng nhiễu ngẫu nhiên

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \delta + \varepsilon_t$$

- Mô Hình Trung Bình Trượt Bậc q - MA(q): trong mô hình trung bình trượt, quá trình được mô tả hoàn toàn bằng tổng có trọng số của các nhiễu ngẫu nhiên hiện hành có độ trễ

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

- Mô Hình Hồi Quy Kết Hợp Trung Bình Trượt - ARIMA(p, d, q)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Nếu xét các chuỗi thời gian thì một chuỗi thời gian có

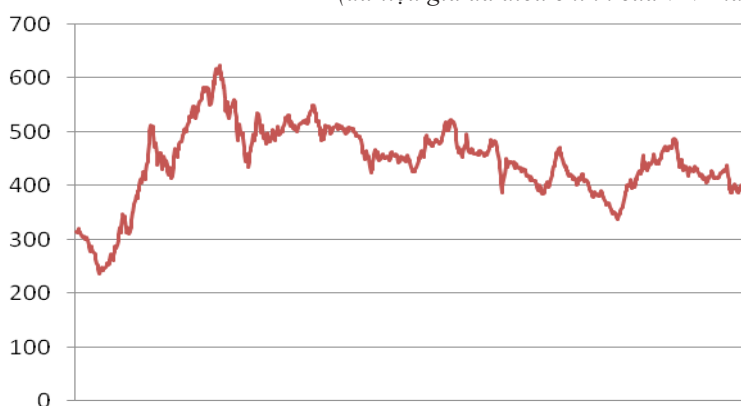
thể dừng hoặc không. Chuỗi sẽ được coi là đồng tích hợp (đồng liên kết) 1 nếu sai phân bậc nhất là chuỗi dừng. Tương ứng chuỗi được gọi là đồng tích hợp bậc d nếu sai phân bậc d là chuỗi dừng. Nếu chuỗi Y_t đồng tích hợp bậc d , ta sẽ có mô hình ARIMA (d, p, q). Trong đó thì d là số lần lấy sai phân chuỗi Y_t để có một chuỗi dừng, p là bậc tự hồi quy, q là bậc trung bình trượt.

Xây dựng mô hình ARIMA cho VN-Index

Nguồn số liệu chỉ số VN-Index được cập nhật liên tục hàng ngày tính từ phiên giao dịch đầu tiên của năm 2009 (ngày 2/1/2009) cho đến ngày 5/10/2012 và đã có sự điều chỉnh số liệu theo các điều chỉnh giá đặc biệt trên thị trường như chia thưởng, chi trả cổ tức,... Đây là giai đoạn thị trường chứng khoán Việt Nam không có quá nhiều sự biến động. Tuy nhiên, VN-Index trong thời gian này phản ánh tương đối tốt tác động của nền kinh tế vĩ mô lên giá chứng khoán.

Biểu đồ 1. Biến động chỉ số VN-Index giai đoạn từ 2/1/2009 đến 5/10/2012

(dữ liệu giá đã điều chỉnh của VN-Index)





Bảng 1. Thống kê dữ liệu VNIndex từ 2/1/2009 đến 5/10/2012

	VNINDEX
Mean	445.5310
Median	450.8500
Maximum	624.1000
Minimum	235.5000
Std. Dev.	69.76274
Observations	924

Như vậy sự biến động giá của chỉ số VN-Index trong giai đoạn này là khá lớn khi giá trị min và max của chuỗi dữ liệu có khoảng cách gần 400 điểm.

Bước 1: Định dạng mô hình

Nhận dạng mô hình ARIMA (p,d,q) là tìm ra giá trị thích hợp của p, d, q với d là bậc sai phân của chuỗi thời gian được khảo sát.

Tuy nhiên, để có được mô hình dự báo tốt, chúng ta sẽ ước lượng các mô hình ARI-MA, MA, AR trên tập số liệu giá từ 2/1/2009- 31/8/2012. Sau đó sẽ dùng mô hình này dự báo cho thời gian từ 1/9/2012- 5/10/2012 để kiểm tra khả năng dự báo của nó.

+ *Kiểm định tính dừng* của chuỗi dữ liệu giá VNIndex từ 2/1/2009- 31/8/2012. Sử dụng kiểm định Dickey- Fuller để kiểm tra tính dừng của chuỗi với giả thiết:

Ho: chuỗi thời gian là không dừng

H1: chuỗi thời gian là dừng

Từ Bảng 2 nhận thấy giá trị thống kê của kiểm định Dickey- Fuller đối với chuỗi VNIndex có Prob. > 0,05 do vậy, bác bỏ Ho, chuỗi VNIndex là không dừng.

Bảng 2. Kết quả kiểm định nghiệm đơn vị xác định tính dừng của chuỗi dữ liệu

Kiểm định		VNIndex	D (VNIndex)
Augmented Dickey-Fuller	ADF Statistic	-2.370956	-23.93014
	Prob.	0.1504	0
Test critical values	1%	-3.437298	-3.437298
	5%	-2.864496	-2.864496
	10%	-2.568397	-2.568397

Tiếp tục kiểm định tính dừng của chuỗi sai phân bậc 1 D (VNIndex) với giả thiết tương tự ta có Prob. < 0,05, bác bỏ giả thiết Ho. Vậy chuỗi sai phân bậc 1 sẽ dừng ở mức ý nghĩa 1%. Như vậy giá trị d=1.

+ *Kiểm định tự tương quan* của chuỗi số liệu

Để định dạng cho mô hình chúng ta sử dụng đồ thị tự tương quan và tự tương quan riêng từng phần của chuỗi d (vnindex). Theo đồ thị trên, tại k=1 có SAC và SPAC đạt cực đại 0,228 và sau đó giảm dần xuống. Thêm vào đó SPAC giảm xuống dưới 5% ở số thứ 2 và giảm dần.

Bước 2: Ước lượng mô hình

Ước lượng các mô hình ARI-MA bằng cách sử dụng tiêu chuẩn Akaike và Schwarz ta có bảng thống kê (Bảng 3).

Theo kết quả từ Bảng 3,

Bảng 3. Kết quả ước lượng mô hình trung bình trượt MA(q) của D(VNIndex)

Mô hình	Akaike info criterion	Schwarz criterion	Mức ý nghĩa
MA(1)	6.729509	6.740061	1%
MA(1)MA(2)	6.729770	6.745597	Không có ý nghĩa

Bảng 4. Thông số kiểm định phần dư từ các mô hình ARIMA

	ARIMA(0,1,1)	ARIMA(1,1,1)	ARIMA(1,1,0)
ADF	-14.42340	-14.46873	-14.52643
Mức ý nghĩa-ADF	1%	1%	1%
Pro(Chi-square(2))	0.0699	0.1508	0.3398

các mô hình ARIMA được đề xuất là ARIMA(0,1,1), ARI-MA(1,1,1) và ARIMA(1,1,0).

Ước lượng 3 mô hình trên để có các thông số để kiểm định tính thích hợp của mô hình bằng cách kiểm định phần dư có được từ các mô hình.

- *Kiểm định tự tương quan, tính dừng của phần dư:*

Dựa trên kiểm định ADF về tính dừng của chuỗi phần dư e tương ứng với từng mô hình ARIMA nhận thấy các e đều là chuỗi dừng ở mức ý nghĩa 1%, các sai số của mô hình đều là nhiễu trắng, và mô hình là phù hợp.

Kiểm định Q với giả thiết Ho là: không có tự tương quan trong chuỗi phần dư. Kiểm định Q cho thấy các phần dư đều không có tự tương quan (sử dụng lược đồ tự tương quan cho các phần dư của từng



mô hình). Sử dụng kiểm định LM về tự tương quan của phần dư với giả thiết:

H₀: e không có tự tương quan.

H₁: e có tự tương quan

Ta có Prob (Chi-square(2)) đều > 0.05. Do vậy các phần dư của các mô hình trên là không có hiện tượng tự tương quan.

Bước 3 : Kiểm định mô hình

So sánh 2 mô hình đã ước lượng ở trên ta có:

Từ Bảng 5 nhận thấy mô

Bảng 5. Kết quả ước lượng mô hình bằng phương pháp OLS

	ARIMA(0,1,1)	ARIMA(1,1,1)	ARIMA(1,1,0)
AIC	6.729509	6.731479	6.730210
Schawrz	6.740061	6.747320	6.740761
HQ	6.733537	6.737526	6.734232

hình mô hình ARIMA(0,1,1) là mô hình phù hợp với dv_{nindex} nhất do có 2 tiêu chuẩn AIC và Schawrz là bé nhất.

Tuy nhiên, để kiểm định mô hình nào là phù hợp nhất, ta sẽ sử dụng thêm tiêu chí lựa chọn mô hình có căn bậc hai của tổng bình phương sai số ngoài mẫu là nhỏ nhất.

Sử dụng phương pháp OLS (bình phương bé nhất) ước lượng với mô hình ARIMA(1,1,0), ARIMA(1,1,1), ARIMA(0,1,1) ta có hệ số của mô hình như sau:

$$dv_{nindex} = 0.089799 + 0.223161MA(1)$$

$$dv_{nindex} = 0.092290 + 0.223161AR(1)$$

$$dv_{nindex} = 0.092162 + 0.125442AR(1) + 0.223161MA(1)$$

Sử dụng 3 mô hình này để dự báo biến động của chỉ số Vnindex trong giai đoạn 1/9/2012

đến 5/10/2012 theo 2 phương trình sau:

$$vnindex_t = 0.089799 + vnindex_{(t-1)} \quad (ARIMA(0,1,1))$$

$$vnindex_t = 0.092290 + vnindex_{(t-1)} \quad (ARIMA(1,1,0))$$

$$vnindex_t = 0.092162 + vnindex_{(t-1)} \quad (ARIMA(1,1,1))$$

Tính toán căn bậc 2 của tổng bình phương sai số ngoài mẫu của 3 chuỗi dữ liệu Vnindex được dự báo từ 1/9/2012 đến 5/10/2012 ta có Bảng 6.

Dựa vào tiêu chí đã nêu trên,

Bảng 6.

Mô hình	Căn bậc hai tổng bình phương sai số ngoài mẫu
ARIMA(0,1,1)	21.67677
ARIMA(1,1,0)	22.43272
ARIMA(1,1,1)	21.67787

mô hình ARIMA(0,1,1) có căn bậc hai tổng bình phương sai số ngoài mẫu là bé nhất. Mô hình ARIMA(0,1,1) là mô hình phù hợp nhất để dự báo chỉ số VNindex.

Dự báo VNIndex bằng mô hình ARIMA(0,1,1)

Sử dụng mô hình trên để dự báo giá VNindex cho 1 phiên giao dịch ngay liền sau 8/10/2012 ta có Vnindex được dự báo đóng cửa ở mức giá 388,23 điểm. So với mức thực tế Vnindex đóng cửa ngày 7/10/2012 là 393,45 điểm. Sai số dự báo so với biến động thực của chỉ số là -1,3%.

Như vậy, mô hình dự báo được chọn có khả năng đưa ra những dự báo gần chính xác về sự biến động giá của Vnindex. Tuy nhiên, do đặc thù của thị trường chứng khoán bị ảnh hưởng bởi rất nhiều các yếu tố khác như thị trường tiền tệ, thị trường bất động sản, hay các thông tin khác có liên quan. Trong khi đó mô hình ARIMA chỉ xem xét dự báo chuỗi thời gian trên cơ sở giá quá khứ. Đây được coi như một nhược điểm làm sai lệch giá trị dự báo của mô hình ARIMA. Như vậy để có được kết quả dự báo chính xác với chuỗi giá thì bản thân giá chứng khoán đã phản ánh đầy đủ các thông tin có thể tác động tới giá hay nói cách khác là thị trường chứng khoán hiệu quả. Đối với thị trường chứng khoán Việt Nam, mức độ hiệu quả của thị trường chưa cao, khả năng phản ánh thông tin của giá còn thấp, do vậy sự sai lệch trong việc dự báo là không thể tránh khỏi ngay cả những phương pháp dự báo khác như phân tích cơ bản hay phân tích kỹ thuật đều có thể gặp phải.

Tóm lại, việc dự báo giá chứng khoán luôn là mối quan tâm lớn của các nhà đầu tư tham gia vào thị trường chứng khoán. Nếu nắm được xu hướng biến động giá chứng khoán nhà đầu tư sẽ có được những quyết định đúng đắn từ đó giảm thiểu được rủi ro khi tham ra thị trường. Mô hình ARIMA chỉ là một trong số các mô hình toán được sử dụng để

xem tiếp trang 10