

# Ứng dụng Mô hình Vector Autoregression (VAR) trong Dự báo Biến động Chỉ số Kinh tế Vĩ mô

Nhóm nghiên cứu

Ngày 8 tháng 12 năm 2025

## Tóm tắt nội dung

**Tóm tắt:** Bài báo này trình bày việc ứng dụng mô hình Vector Autoregression (VAR) để dự báo chỉ số giá tiêu dùng cơ bản (Core CPI). Nghiên cứu sử dụng bộ dữ liệu kinh tế vĩ mô bao gồm kỳ vọng lạm phát, tín dụng tư nhân và chỉ số CPI tổng thể. Kết quả thực nghiệm cho thấy mô hình VAR với độ trễ tối ưu  $p = 3$  đạt được  $R^2$  khoảng 0.19 và sai số phần trăm trung bình tuyệt đối (MAPE) là 14.07%.

**Keywords:** VAR, Forecasting, Inflation, Core CPI, Time Series. **JEL codes:** C32, E31, E37.

## 1 Introduction (Giới thiệu)

Trong bối cảnh kinh tế hiện đại, việc dự báo chính xác lạm phát là yếu tố then chốt để hoạch định chính sách tiền tệ. Bài báo này tập trung giải quyết bài toán dự báo biến phụ thuộc **Core CPI** dựa trên sự tương tác đa chiều giữa các biến kinh tế vĩ mô. Mục tiêu chính là xây dựng một mô hình VAR ổn định, đánh giá hiệu quả dự báo thông qua các chỉ số sai số và đóng góp một góc nhìn định lượng về mối quan hệ giữa các biến số tài chính - tiền tệ với lạm phát.

## 2 Literature Review (Tổng quan nghiên cứu)

Máy học (Machine Learning) và các mô hình kinh tế lượng truyền thống như VAR đã được sử dụng rộng rãi trong kinh tế vĩ mô. Các nghiên cứu trước đây (Sims, 1980) đã chỉ ra rằng VAR có ưu thế trong việc nắm bắt sự phụ thuộc lẫn nhau giữa các chuỗi thời gian mà không cần đặt ra các giả định quá chặt chẽ về cấu trúc nhân quả một chiều.

## 3 Methodology (Phương pháp nghiên cứu)

### 3.1 Mô hình Vector Autoregression (VAR)

Mô hình VAR mở rộng mô hình tự hồi quy đơn biến (AR) sang trường hợp đa biến vector. Cấu trúc toán học của mô hình VAR bậc  $p$  được biểu diễn như sau:

$$Y_t = c + A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + \epsilon_t \quad (1)$$

Trong đó:  $Y_t$  là vector ( $k \times 1$ ) của các biến nội sinh;  $c$  là vector hằng số;  $A_i$  là các ma trận hệ số và  $\epsilon_t$  là nhiễu trắng.

## 3.2 Quy trình và Metrics đánh giá

Các thước đo hiệu quả (Metrics) được sử dụng bao gồm: RMSE (Root Mean Squared Error), MAPE (Mean Absolute Percentage Error) và R-squared ( $R^2$ ).

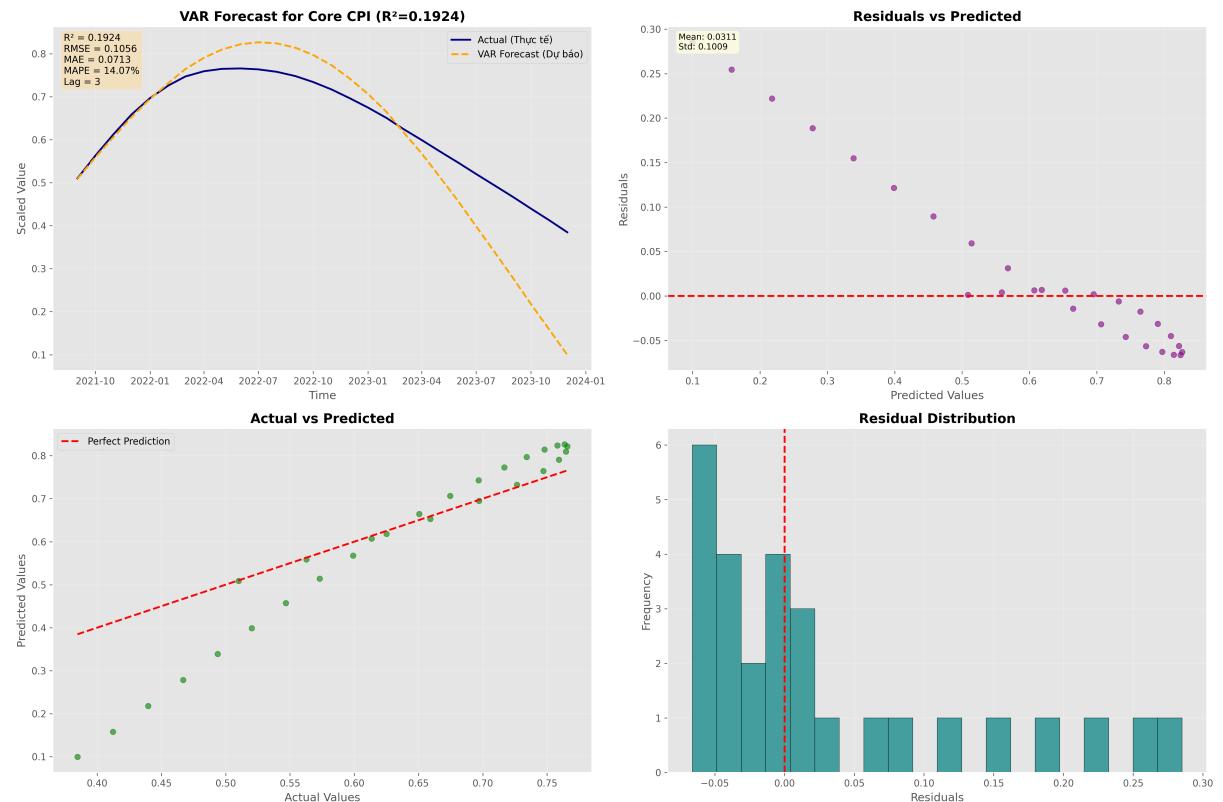
## 4 Data (Dữ liệu)

### 4.1 Nguồn dữ liệu và Xử lý

Dữ liệu được chia thành tập huấn luyện (train) và tập kiểm tra (test) dựa trên chuỗi thời gian. Các biến đầu vào được sử dụng ở dạng trễ bậc 1 ('lag\_1') để làm biến đại diện (proxy).

### 4.2 Lựa chọn biến (Feature Selection)

Dựa trên các phân tích thống kê và chẩn đoán mô hình dưới đây, chúng tôi đã xác định được các biến phù hợp để đưa vào mô hình.



Hình 1: Kết quả chẩn đoán biến và mô hình (Diagnostics)

4 biến quan trọng nhất bao gồm: Core CPI (Mục tiêu), Kỳ vọng lạm phát 12 tháng, Tín dụng tư nhân, và CPI tổng thể.

## 5 Results (Kết quả)

### 5.1 Thống kê hiệu quả mô hình (Model Metrics)

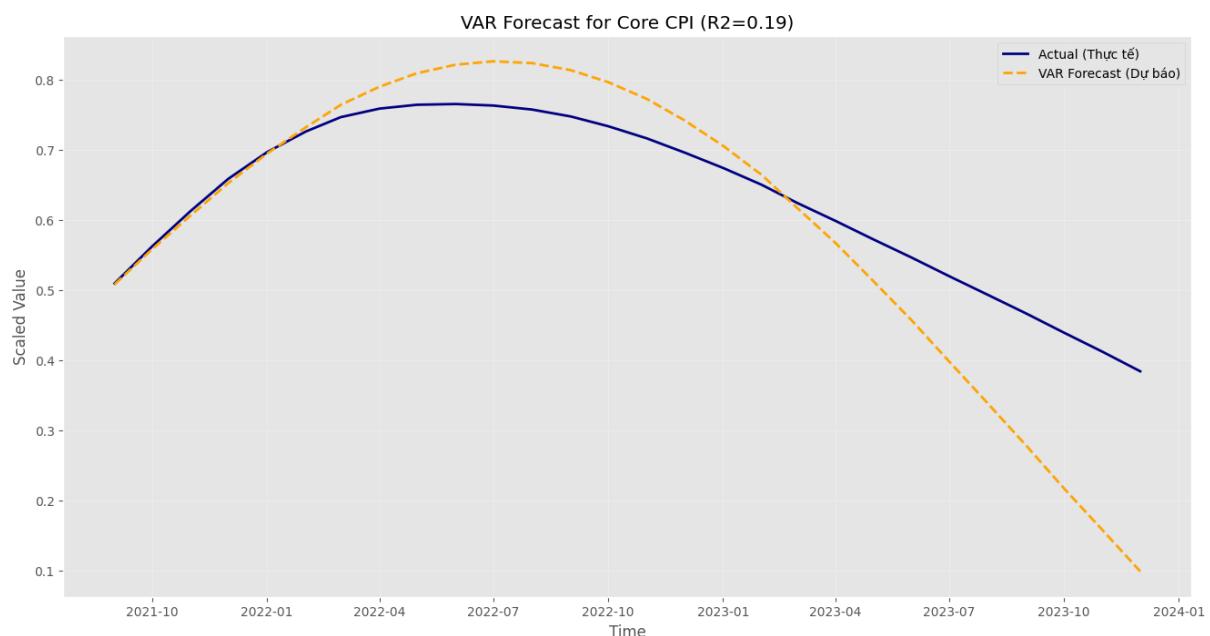
Kết quả đánh giá định lượng trên tập kiểm tra (Test set) được tổng hợp trong Bảng 1.

Bảng 1: Các chỉ số đánh giá hiệu quả mô hình VAR

Metric	Giá trị
Optimal Lag (AIC)	3.00
R-squared ( $R^2$ )	0.1924
RMSE	0.1056
MAE	0.0713
MAPE	14.07 %

### 5.2 Kết quả dự báo chi tiết

Biểu đồ dưới đây minh họa sự so sánh giữa giá trị thực tế (Actual) và giá trị dự báo (Predicted) của lạm phát cơ bản.



Hình 2: Biểu đồ so sánh Core CPI Thực tế và Dự báo

Chi tiết dữ liệu dự báo cho 10 kỳ gần nhất được trình bày trong Bảng 2.

Bảng 2: Dữ liệu chi tiết: Thực tế vs Dự báo (10 kỳ gần nhất)

Date	Actual	Predicted	Difference
2024-05-01	0.4993	0.6034	-0.1041
2024-06-01	0.5282	0.5051	+0.0231
2024-07-01	0.4891	0.5733	-0.0842
2024-08-01	0.5488	0.4705	+0.0783
2024-09-01	0.2848	0.4285	-0.1437
2024-10-01	0.3646	0.3013	+0.0633
2024-11-01	0.3957	0.3470	+0.0487
2024-12-01	0.3708	0.4827	-0.1119
2025-01-01	0.1802	0.3005	-0.1203
2025-02-01	0.3725	0.2319	+0.1406

\*Difference = Actual - Predicted

## 6 Conclusions (Kết luận)

Nghiên cứu đã xây dựng thành công framework dự báo Core CPI bằng mô hình VAR với độ trễ tối ưu là 3. Mặc dù RMSE thấp (0.1056), chỉ số  $R^2$  (0.1924) cho thấy cần mở rộng thêm các biến vĩ mô khác hoặc sử dụng mô hình phi tuyến để cải thiện độ chính xác.

## References (Tài liệu tham khảo)

- Cabrera Bonilla. (2024). *Tiêu đề bài báo gốc...*
- Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 48(1), 1-48.
- Statsmodels Developer. (2024). Statsmodels: Statistical models, hypothesis tests, and data exploration.

## Appendix (Phụ lục)

### Danh sách biến và Siêu tham số:

- **Thư viện sử dụng:** statsmodels.tsa.api.VAR
- **Lag order (Max):** 3
- **Hàm măt mát:** Least Squares (OLS).