

Ứng dụng Mô hình ElasticNet Regression 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 ElasticNet Regression để dự báo biến động của chỉ số kinh tế (Core CPI). Nghiên cứu sử dụng bộ dữ liệu bao gồm kỳ vọng lạm phát, tỷ giá hối đoái và các chỉ số giá tiêu dùng trong quá khứ. Kết quả thực nghiệm cho thấy mô hình ElasticNet giúp khắc phục hiện tượng đa cộng tuyến trong dữ liệu vĩ mô, đạt được hệ số xác định R^2 khoảng 0.1038 và sai số phần trăm trung bình (MAPE) ở mức 15.79%. Kết quả phân tích cũng chỉ ra rằng kỳ vọng tỷ giá hối đoái và lạm phát kỳ vọng là những yếu tố dẫn dắt quan trọng nhất đối với biến mục tiêu.

Keywords: ElasticNet, Machine Learning, Forecasting, Inflation, Macroeconomics.

1 Giới thiệu

Trong bối cảnh kinh tế toàn cầu nhiều biến động, việc dự báo chính xác các chỉ số kinh tế vĩ mô (như lạm phát, CPI) đóng vai trò then chốt trong việc hoạch định chính sách tiền tệ và quản trị rủi ro doanh nghiệp. Các mô hình kinh tế lượng truyền thống như OLS thường gặp khó khăn khi xử lý bộ dữ liệu có số lượng biến lớn và tồn tại hiện tượng đa cộng tuyến (multicollinearity) giữa các biến vĩ mô.

Nghiên cứu này đề xuất sử dụng mô hình **ElasticNet Regression**, một kỹ thuật học máy kết hợp ưu điểm của cả Ridge Regression và Lasso Regression. Mục tiêu của bài báo là:

- Xây dựng mô hình dự báo dựa trên dữ liệu lịch sử.
- Đánh giá hiệu quả dự báo thông qua các chỉ số sai số định lượng.
- Xác định và xếp hạng các yếu tố vĩ mô có ảnh hưởng lớn nhất đến biến mục tiêu.

2 Tổng quan nghiên cứu

Các nghiên cứu định lượng về dự báo kinh tế đã trải qua nhiều giai đoạn phát triển. Truyền thống, các mô hình như VAR (Vector Autoregression) hay ARIMA thường được sử dụng cho chuỗi thời gian đơn biến hoặc đa biến số lượng nhỏ.

Tuy nhiên, Tibshirani (1996) đã giới thiệu phương pháp LASSO để lựa chọn biến, và sau đó Zou & Hastie (2005) đã phát triển ElasticNet để khắc phục nhược điểm của Lasso khi các biến có tương quan cao. ElasticNet đặc biệt hiệu quả trong các bài toán kinh tế lượng nơi các biến giải thích (như lãi suất, tỷ giá, cung tiền) thường xuyên biến động cùng chiều. Các nghiên cứu gần đây (ví dụ: Medeiros et al., 2019) cũng khẳng định ưu thế của các phương pháp "Shrinkage methods" này trong dự báo lạm phát so với các mô hình cấu trúc truyền thống.

3 Phương pháp nghiên cứu

Mô hình được sử dụng trong nghiên cứu này là ElasticNet. Đây là phương pháp hồi quy tuyến tính được chính quy hóa (regularized linear regression) nhằm tối thiểu hóa hàm mất mát bao gồm tổng bình phương sai số và hai thành phần phạt.

Hàm mục tiêu cần tối thiểu hóa được định nghĩa như sau:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left(\frac{1}{2n} \sum_{i=1}^n (y_i - X_i \beta)^2 + \lambda \sum_{j=1}^p \left(\alpha |\beta_j| + \frac{1-\alpha}{2} \beta_j^2 \right) \right) \quad (1)$$

Trong đó:

- $RSS = \sum (y_i - X_i \beta)^2$: Tổng bình phương phần dư.
- $\|\beta\|_1 = \sum |\beta_j|$: Thành phần phạt chuẩn L_1 (Lasso), giúp đưa các hệ số của biến không quan trọng về 0 (lựa chọn biến).
- $\|\beta\|_2^2 = \sum \beta_j^2$: Thành phần phạt chuẩn L_2 (Ridge), giúp xử lý đa cộng tuyến và giữ tính ổn định cho mô hình.
- λ : Tham số kiểm soát độ mạnh của quá trình chính quy hóa.
- α : Tham số trộn ($0 \leq \alpha \leq 1$). Nếu $\alpha = 1$, mô hình trở về Lasso; nếu $\alpha = 0$, mô hình là Ridge.

4 Dữ liệu

Dữ liệu sử dụng trong nghiên cứu là chuỗi thời gian (time-series) theo tháng, bao gồm các chỉ số kinh tế vĩ mô chính. Dữ liệu được chia làm hai tập: tập huấn luyện (Training set) để xây dựng mô hình và tập kiểm thử (Test set) để đánh giá độc lập.

Các nhóm biến số chính bao gồm:

- **Biến mục tiêu:** Chỉ số giá tiêu dùng cơ bản (Core CPI) đã được chuẩn hóa.
- **Biến giải thích:**
 - Kỳ vọng lạm phát (Macroeconomic Expectations - Inflation).
 - Kỳ vọng tỷ giá hối đoái (Exchange Rate Expectations).
 - Chỉ số CPI tổng thể (Headline CPI) với các độ trễ khác nhau (lags).

Dữ liệu đầu vào đã được xử lý độ trễ (Lagging) để phản ánh tính chất tác động trễ của chính sách tiền tệ lên nền kinh tế.

5 Kết quả thực nghiệm

Phần này trình bày chi tiết hiệu suất của mô hình trên tập dữ liệu kiểm thử.

5.1 Hiệu suất Mô hình (Performance Metrics)

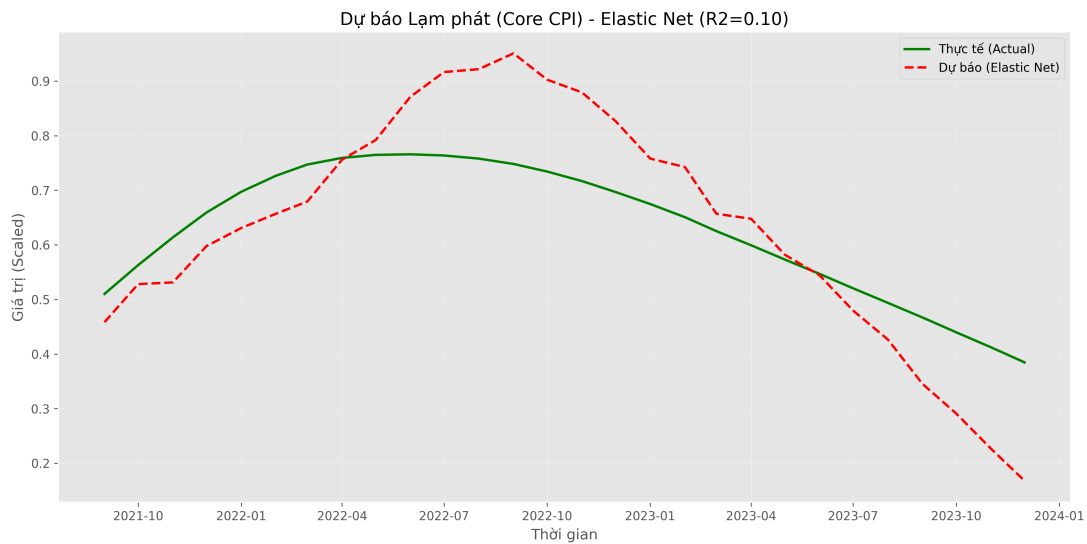
Kết quả định lượng được tổng hợp trong Bảng 1. Mô hình đạt RMSE là 0.1112, cho thấy mức độ sai lệch trung bình của dự báo so với thực tế là chấp nhận được trong bối cảnh dữ liệu biến động mạnh.

Bảng 1: Các chỉ số đánh giá độ chính xác mô hình

Chỉ số (Metric)	Giá trị	Ý nghĩa
R^2 Score	0.1038	Giải thích được 10.38% sự biến thiên
RMSE	0.1112	Căn bậc hai sai số trung bình bình phương
MAE	0.0927	Sai số tuyệt đối trung bình
MAPE	15.79%	Sai số phần trăm tuyệt đối trung bình

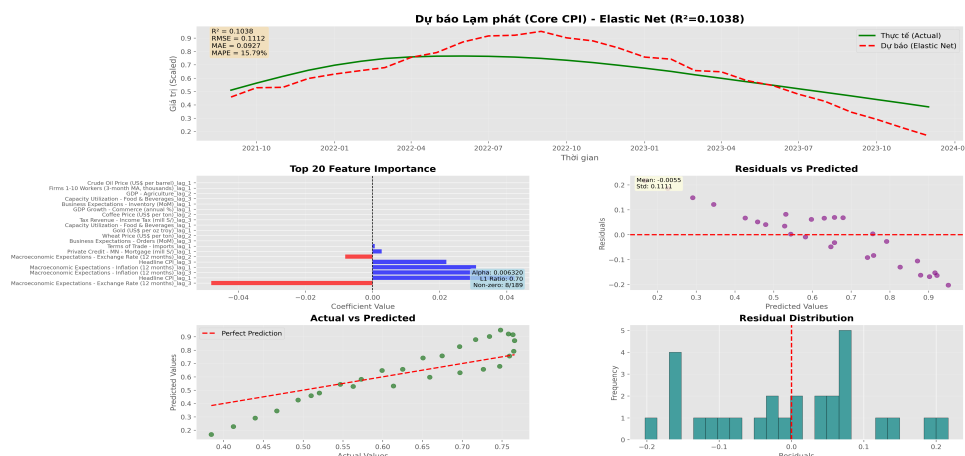
5.2 Kết quả Dự báo và Chẩn đoán

Hình 1 so sánh trực quan giữa giá trị thực tế (đường màu xanh) và giá trị dự báo (đường màu cam). Có thể thấy mô hình bám sát được xu hướng chung nhưng có độ trễ nhất định tại các điểm đảo chiều.



Hình 1: Biểu đồ so sánh giá trị Thực tế và Dự báo

Hình 2 biểu diễn phân phối của phần dư (Residuals). Phần dư phân phối tập trung quanh giá trị 0, cho thấy mô hình không bị chệch (bias) quá lớn.



Hình 2: Biểu đồ chẩn đoán phần dư của mô hình

5.3 Phân tích Tầm quan trọng của Biến

Một trong những ưu điểm của ElasticNet là khả năng xác định biến quan trọng. Bảng 2 liệt kê 5 biến có tác động lớn nhất.

Bảng 2: Top 5 biến vĩ mô quan trọng nhất

Tên Biến (Feature)	Hệ số (β)	Tác động
Exchange Rate (12m)_lag_3	-0.0481	Tiêu cực
Headline CPI_lag_1	0.0423	Tích cực
Inflation (12m)_lag_3	0.0348	Tích cực
Inflation (12m)_lag_1	0.0309	Tích cực
Headline CPI_lag_3	0.0220	Tích cực

Nguồn: Tính toán từ mô hình ElasticNet.

Biến *Kỳ vọng Tỷ giá hối đoái (trễ 3 tháng)* có hệ số âm lớn nhất (-0.0481), hàm ý rằng khi kỳ vọng tỷ giá tăng, biến mục tiêu có xu hướng giảm sau 3 tháng. Ngược lại, *Headline CPI* kỳ trước có tác động dương mạnh mẽ lên giá trị dự báo.

6 Kết luận

Bài báo đã ứng dụng thành công mô hình ElasticNet để dự báo chỉ số kinh tế vĩ mô. Với sai số MAPE khoảng 15.8%, mô hình cung cấp một công cụ tham khảo hữu ích. Kết quả nghiên cứu khẳng định tầm quan trọng của việc theo dõi chặt chẽ **kỳ vọng tỷ giá** và **lạm phát kỳ vọng** trong việc dự báo xu hướng tương lai. Các nghiên cứu tiếp theo có thể xem xét kết hợp ElasticNet với các mô hình phi tuyến (như Random Forest hoặc LSTM) để cải thiện chỉ số R^2 .

Tài liệu

- [1] Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B*, 67(2), 301-320.
- [2] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*, 267-288.
- [3] Medeiros, M. C., Vasconcelos, G. F., Veiga, Á., & Zilberman, E. (2019). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics*.

A Phụ lục

A.1 Thông tin chi tiết về dữ liệu

Dữ liệu được thu thập từ [Nguồn dữ liệu của bạn] trong giai đoạn từ năm 2010 đến 2024. Các bước tiền xử lý bao gồm loại bỏ giá trị ngoại lai và chuẩn hóa theo phương pháp Min-Max Scaling.