

KHÁM PHÁ

ĐỊNH LUẬT OKUN VỚI PHƯƠNG PHÁP THỐNG KÊ HIỆN ĐẠI

TS. Phạm Đăng Quyết*

Giới thiệu

Khi nói đến nghiên cứu nền kinh tế, tăng trưởng và việc làm là hai yếu tố chính mà các nhà kinh tế phải xem xét. Có một mối quan hệ rõ ràng giữa tăng trưởng và việc làm, và nhiều nhà kinh tế đã đóng góp vào cuộc thảo luận bằng cách cố gắng nghiên cứu mối quan hệ giữa tăng trưởng kinh tế và mức thất nghiệp.

Nhà kinh tế học Arthur Okun lần đầu tiên bắt đầu giải quyết cuộc thảo luận như vậy vào những năm 1960 và nghiên cứu của ông về chủ đề này đã được gọi là định luật Okun. Định luật này cho biết mối quan hệ giữa tỷ lệ thất nghiệp và sự thay đổi của sản lượng trong một quốc gia. Cơ quan nghiên cứu kinh tế của Ngân hàng Dự trữ Liên bang St. Louis giải thích rằng định luật Okun "nhằm cho chúng ta biết tổng sản phẩm trong nước (GDP) của một quốc gia có thể bị mất bao nhiêu khi tỷ lệ thất nghiệp cao hơn tỷ lệ tự nhiên" [3]. Logic đằng sau định luật Okun rất đơn giản: Sản lượng phụ thuộc vào lượng lao động được sử dụng trong quá trình sản xuất, do đó có mối quan hệ thuận giữa sản lượng và việc làm. Tổng số việc làm bằng lực lượng

lao động trừ đi số thất nghiệp, do đó có mối quan hệ nghịch giữa sản lượng và thất nghiệp (có điều kiện về lực lượng lao động).

1. Tổng quan

Theo định luật này, khi tăng trưởng kinh tế tăng lên, tỷ lệ thất nghiệp sẽ giảm và ngược lại, khi tăng trưởng kinh tế giảm, tỷ lệ thất nghiệp sẽ tăng lên. Hình 1 biểu diễn tốc độ tăng trưởng và tỷ lệ thất nghiệp theo quý từ năm 2011 đến 2023 ở Việt Nam.

Định luật Okun có thể được biểu diễn bằng công thức sau [1]:

$$\frac{Y - \bar{Y}}{\bar{Y}} = -c(u - \bar{u}) \quad (1)$$

Ở đây: Y là sản lượng thực tế, \bar{Y} là sản lượng tiềm năng, u là tỉ lệ thất nghiệp thực tế, \bar{u} là tỉ lệ thất nghiệp tự nhiên, c là hệ số phản ánh mối quan hệ giữa thay đổi của tỷ lệ thất nghiệp và thay đổi của sản lượng.

Công thức (1) trên rất khó ứng dụng vì \bar{Y} và \bar{u} chỉ có thể ước lượng chứ không thể thống kê hay đo lường. Một dạng phổ biến hơn của Định luật Okun, sử dụng mức thay đổi của sản lượng so với mức thay đổi của thất nghiệp, dưới dạng như sau:

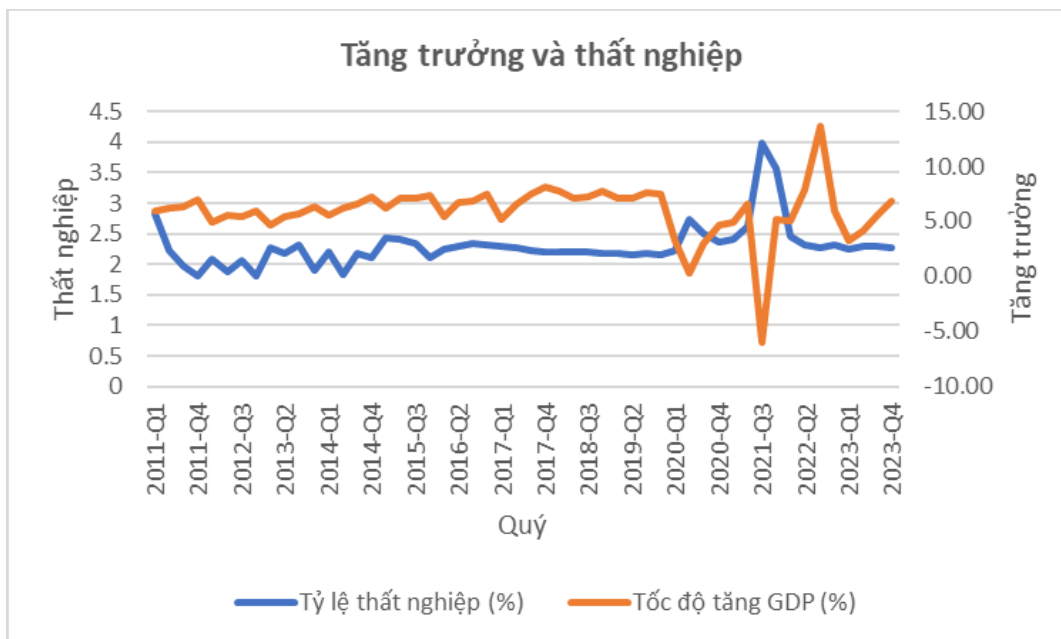
$$\frac{\Delta Y}{Y} = k - c\Delta u \quad (2)$$

* Hội Thống kê Việt Nam

➤➤➤ NGHIÊN CỨU • TRAO ĐỔI

Trong đó: Y và c như định nghĩa ở trên; ΔY là mức thay đổi của sản lượng thực tế; Δu là mức thay đổi của tỷ lệ thất nghiệp thực tế; k là tốc độ tăng trưởng bình quân theo năm của sản lượng toàn dụng; $(\frac{\Delta Y}{Y})$ chính là tốc độ tăng của sản lượng Y .

Hình 1. Tốc độ tăng trưởng GDP và tỷ lệ thất nghiệp theo quý giai đoạn 2011-2023



Nguồn: Tổng cục Thống kê

Điều quan trọng cần lưu ý, định luật Okun là một mối quan hệ thống kê dựa trên sự hồi quy của tỷ lệ thất nghiệp và tăng trưởng kinh tế. Định luật Okun là một công cụ hữu ích trong việc đánh giá, dự báo xu hướng giữa tỷ lệ thất nghiệp và sản lượng GDP thực tế.

Như vậy, chạy hồi quy có thể dẫn đến các hệ số khác nhau được sử dụng để giải quyết sự thay đổi thất nghiệp, dựa trên cách nền kinh tế phát triển. Tất cả phụ thuộc vào khoảng thời gian được sử dụng và đầu vào, đó là dữ liệu GDP và tỷ lệ thất nghiệp lịch sử.

Chúng ta hãy xem xét mối tương quan giữa tăng trưởng hàng quý về GDP thực tế so với những thay đổi hàng quý về tỷ lệ thất nghiệp từ nguồn dữ liệu sẵn có của Tổng cục Thống kê giai đoạn 2011-2023.

2. Các phương pháp

2.1. Hồi quy thống kê truyền thống

Mục tiêu của mô hình hồi quy tuyến tính truyền thống là tìm ra một phương trình có dạng $y = ax + b$, trong đó a và b là các hệ số cố định, x là biến độc lập và y là biến phụ thuộc. Phương trình này cho phép dự đoán giá trị của y khi biết giá trị của x , hoặc ngược lại. Định luật Okun thể hiện mối tương quan tuyến tính giữa một biến phụ thuộc là tốc độ tăng GDP và một biến độc lập là mức thay đổi tỷ lệ thất nghiệp (Hình 2).

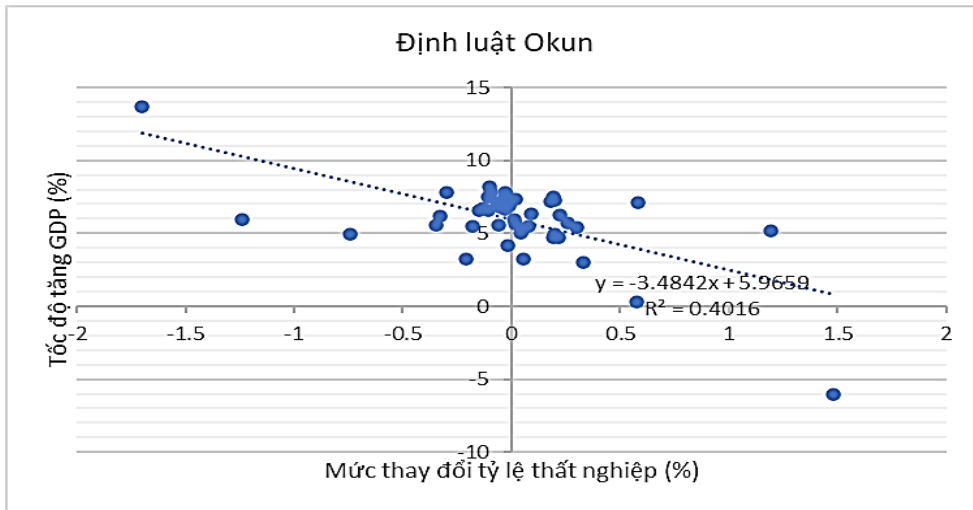
Mô hình hồi quy trên cho biết mối quan hệ thực nghiệm giữa tăng trưởng kinh tế và mức thất nghiệp, theo đó tỷ lệ thất nghiệp theo quý tăng lên 1% khiến tăng trưởng GDP theo quý giảm 3,5%, hay ngược lại, tỷ lệ thất

nghiệp giảm 1% khiến tăng trưởng GDP tăng lên 3,5%.

Ở đây hệ số xác định $R^2 = 0,402$. Điều này có nghĩa là mô hình hồi quy tuyến tính của chúng ta giải thích được khoảng 40,2%

phương sai của biến phụ thuộc, hay nói cách khác, định luật Okun không phải là một mối quan hệ chặt chẽ, nhưng nó dự đoán rằng sự suy giảm tăng trưởng thường trùng với tỷ lệ thất nghiệp gia tăng.

Hình 2. Mối quan hệ giữa tốc độ tăng GDP và mức độ thay đổi tỷ lệ thất nghiệp hàng quý giai đoạn 2012-2023



Với tiến bộ công nghệ, các phương pháp thống kê hiện đại đã được phát triển trong những thập kỷ gần đây với việc sử dụng các công cụ toán học, máy tính và lý thuyết xác suất nâng cao để giải quyết các bài toán thống kê phức tạp và đa dạng. Chúng ta sẽ sử dụng một số phương pháp hồi quy thống kê hiện đại (Máy học và Bayes) để xem xét lựa chọn mô hình hồi quy nào là thích hợp hơn. Chúng tôi sử dụng tiêu chuẩn R^2 để đánh giá hiệu suất và lựa chọn mô hình hồi quy tối ưu. Nếu hệ số xác định có giá trị lớn hơn 50% thì có thể coi mô hình khá phù hợp với tập dữ liệu. Còn nếu hệ số xác định nhỏ hơn 30% thì có thể khẳng định mô hình không phù hợp và nên tìm mô hình khác. Mô hình tốt là một mô hình có khả năng cung cấp các kết quả dự báo càng sát với thực tế càng tốt.

2.2. Hồi quy máy học

Hồi quy máy học có thể được coi là một trường hợp riêng biệt của hồi quy thống kê, nhưng với sự tham gia của các thuật toán máy học để tối ưu hóa các tham số của mô hình hồi quy. Chúng tôi sử dụng thư viện Keras để xây dựng mô hình hồi quy. Mạng nơ-ron cho hồi quy là một loại mô hình thống kê sử dụng trọng số thích ứng và có thể xấp xỉ các hàm phi tuyến tính của đầu vào của nó. Nó là một mạng lưới thần kinh được đào tạo, chuyển tiếp và được kết nối đầy đủ có thể được sử dụng để hồi quy, dự đoán và phân loại, bao gồm ba thành phần chính:

Lớp đầu vào, là nơi các quan sát huấn luyện được cung cấp. Số lượng biến tiên lượng cũng được chỉ định ở đây thông qua các tế bào thần kinh.

➤➤➤ NGHIÊN CỨU • TRAO ĐỔI

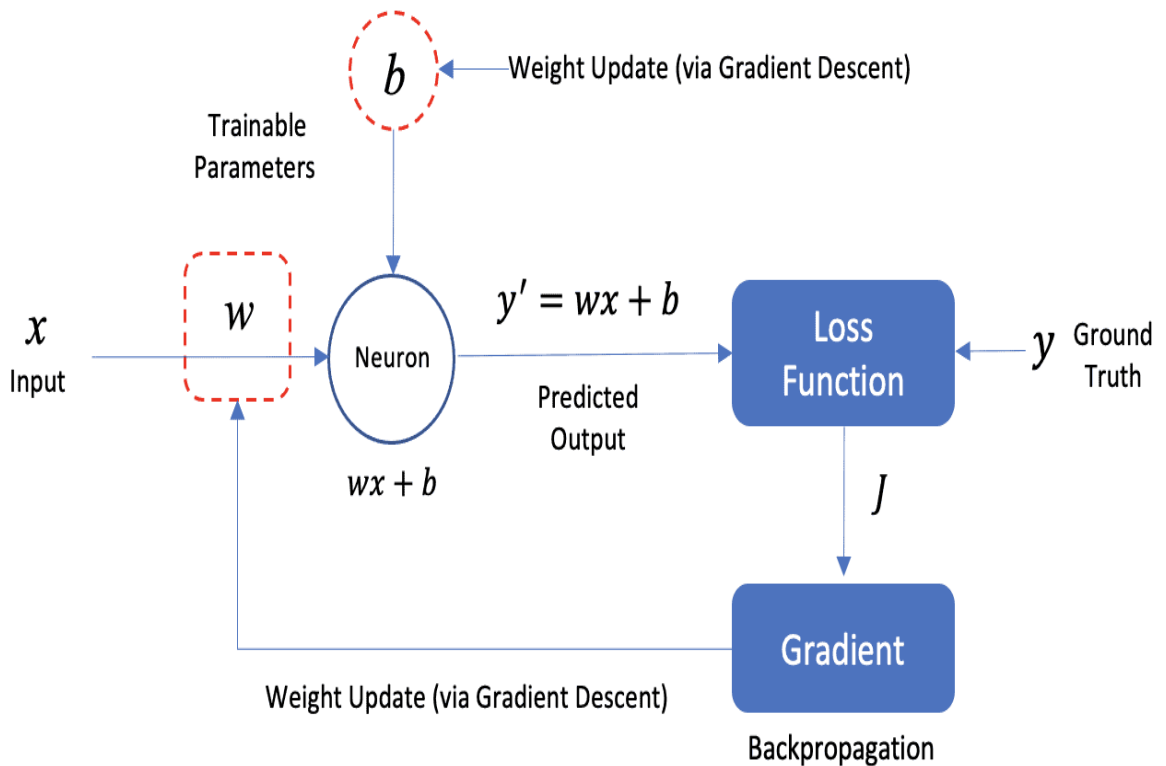
Lớp ẩn, là các lớp trung gian giữa các lớp đầu vào và đầu ra. Mạng nơ-ron sâu tìm hiểu về các mối quan hệ liên quan đến dữ liệu trong thành phần này.

Lớp đầu ra, là lớp mà đầu ra cuối cùng được trích xuất từ những gì đang xảy ra trong hai lớp trước. Trong trường hợp có vấn đề hồi quy, đầu ra sau đó sẽ có một tế bào thần kinh.

Một sơ đồ chính thức hơn một chút được hiển thị bên dưới cho thấy cách các

tham số mô hình (w và b) được cập nhật trong quá trình huấn luyện. Các tham số mô hình được khởi tạo thành các giá trị ngẫu nhiên nhỏ. Trong quá trình huấn luyện, khi dữ liệu huấn luyện được truyền qua mạng, giá trị dự đoán của mô hình (y') được so sánh với giá trị thực (y) cho một mẫu nhất định từ tập dữ liệu. Các giá trị này được sử dụng làm cơ sở để tính toán tổn thất, sau đó được sử dụng làm phản hồi trong mạng để điều chỉnh các tham số mô hình theo cách cải thiện dự đoán [4].

Hình 3. Sơ đồ mạng nơ-ron của Hồi quy Keras



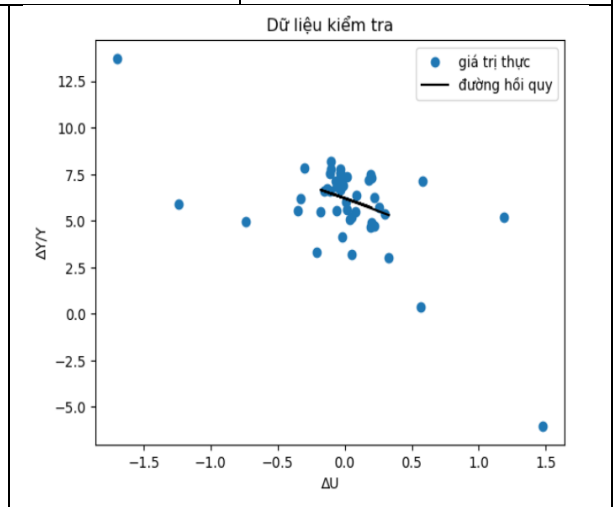
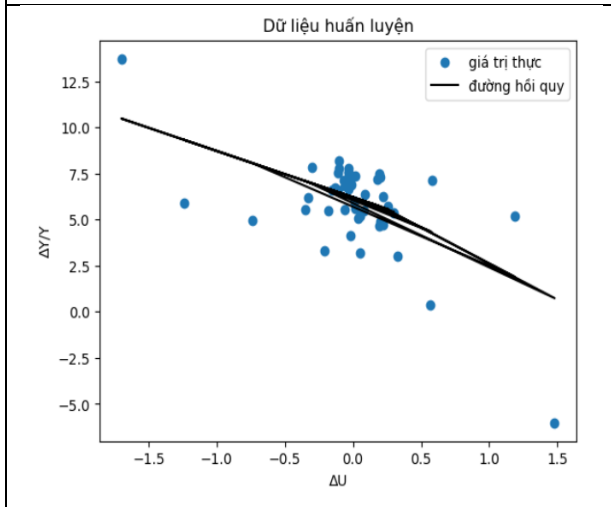
Quá trình cập nhật trọng số này bao gồm hai bước được gọi là Gradient Descent và Backpropagation. Ở giai đoạn này, điều quan trọng là phải hiểu các chi tiết toán học về cách các thuật toán này hoạt động, nhưng

điều quan trọng nữa là phải hiểu rằng có một quá trình lặp đi lặp lại để huấn luyện mô hình.

Kết quả nhận được sau khi chạy hồi quy Keras như sau:

Hình 4. Mô hình hồi quy Keras với bộ dữ liệu huấn luyện và kiểm tra

Model: "sequential"			Hệ số hồi quy: Intercept (k): 0.342 Slope (c): -0.176 Hệ số xác định: R^2 : 0.430
Layer (type)	Output Shape	Param #	
dense (Dense)	(None, 32)	64	
dense_1 (Dense)	(None, 16)	528	
dense_2 (Dense)	(None, 1)	17	
===== Total params: 609 (2.38 KB) Trainable params: 609 (2.38 KB) Non-trainable params: 0 (0.00 Byte)			



Hệ số xác định R^2 của mô hình Keras là 43% lớn hơn hệ số này của mô hình hồi quy thống kê truyền thống là 40,2%. Điều đó cho thấy mô hình hồi quy máy học do được huấn luyện nên phù hợp hơn.

Hệ số hồi quy $c = -0,18$ cho biết tỷ lệ thất nghiệp tăng lên mỗi 1 điểm phần trăm sẽ dẫn đến tăng trưởng GDP thực giảm 0,18%.

Các mô hình hồi quy thống kê truyền thống hay máy học trên thường được coi là cách tiếp cận thống kê tần suất, là một phương pháp thống kê giúp chúng ta biểu diễn và phân tích dữ liệu theo số lần xuất hiện của các giá trị hoặc nhóm giá trị trong

tập dữ liệu. Thống kê tần suất chỉ đơn giản lấy xác suất của một sự kiện nhất định dựa trên các bộ dữ liệu quan sát, kết quả đầu ra là một ước tính điểm, các tham số được xem là không biết luôn không đổi, còn các biến dự đoán là ngẫu nhiên.

Bây giờ chúng ta sẽ tiếp cận mô hình hồi quy theo một cách tiếp cận khác, tiếp cận Bayes, theo đó các yếu tố dự đoán được coi là không đổi trong khi các ước tính tham số là ngẫu nhiên và mỗi tham số tuân theo một phân bố với một giá trị trung bình và phương sai, lúc này kết quả đầu ra không còn là một ước tính điểm đơn lẻ nữa mà là một phân bố.

➤➤➤ NGHIÊN CỨU • TRAO ĐỔI

2.3. Hồi quy Bayes

Hồi quy Bayes là một phương pháp hồi quy dựa trên nguyên lý Bayes, trong đó các tham số của mô hình hồi quy được xem như là các biến ngẫu nhiên có phân bố xác suất. Phương pháp này cho phép ta tính toán phân bố hậu nghiệm (posterior) của các tham số dựa trên phân bố tiên nghiệm (prior) và dữ liệu quan sát. Công thức định lý Bayes có dạng:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Trong đó:

- $P(A|B)$ là xác suất hậu nghiệm của sự kiện A khi biết sự kiện B đã xảy ra.
- $P(B|A)$ là xác suất có điều kiện của sự kiện B khi biết sự kiện A đã xảy ra, còn gọi là khả năng của B.
- $P(A)$ là xác suất tiên nghiệm của sự kiện A, không quan tâm đến sự kiện B.

```
Mean Intercept (k):      coef: 5.9646 std err: 0.291
Mean slope (c):         coef: -3.4863 std err: 0.627
Mean MSE (sigma):      3.888
```

Tham số tiên nghiệm của mô hình hồi quy Bayes được ước tính như sau:

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
Intercept	5.96	0.20	5.58	6.34	0.0	0.0	10551.41	6223.15	1.0
slope	-3.49	0.45	-4.38	-2.70	0.0	0.0	12397.40	6416.65	1.0
sigma	2.04	0.22	1.64	2.45	0.0	0.0	11414.24	5215.05	1.0

Đối với mỗi tham số, `mcse_mean` và `mcse_sd` là trung bình và sai số chuẩn Monte Carlo, `ess_bulk` là sai số bình phương trung bình (MSE), `ess_tail` là hệ số R^2 của biến và `r_hat` là hệ số giảm quy mô tiềm

• $P(B)$ là xác suất biên của sự kiện B, còn gọi là hằng số chuẩn hóa.

Chúng ta sẽ bắt đầu với hồi quy tuyến tính $Y = X\beta + \epsilon$

Cú pháp chuẩn cho hồi quy tuyến tính Bayes được đưa ra như sau [2]:

$$y \sim N(X\beta, \sigma^2)$$

Ở đây, biến kết quả không còn là ước tính điểm nữa mà là phân bố chuẩn với giá trị trung bình của $X\beta$ và phương sai σ^2 , trong đó $X\beta$ là phương trình tuyến tính tổng quát trong X.

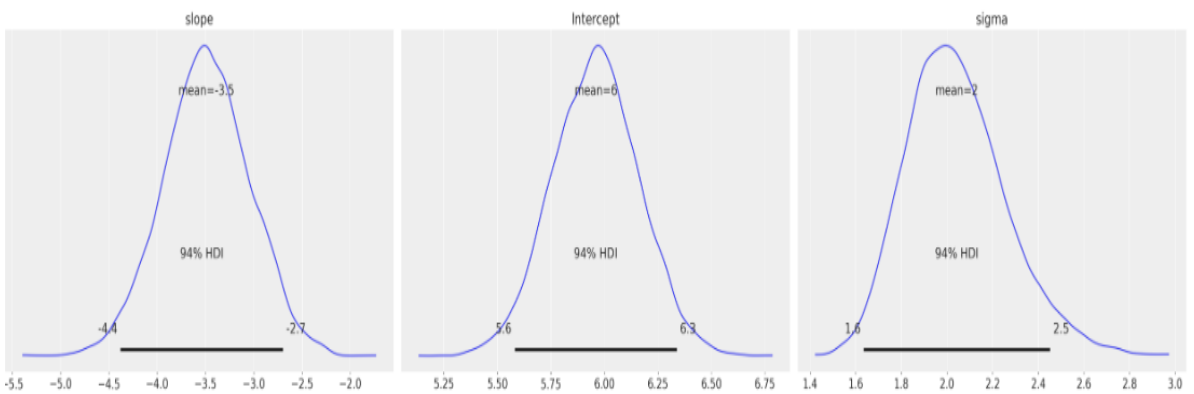
Mục đích của chúng ta là cập nhật phân bố của các tham số chưa biết β và σ^2 dựa trên dữ liệu X quan sát được.

Nếu chúng ta coi các thông tin tiên nghiệm của mô hình Bayes là các tham số của mô hình hồi quy tuyến tính truyền thống, ta có các thông tin tiên nghiệm như sau:

năng trên các chuỗi phân tách (tại độ hội tụ `r_hat=1`).

Chúng ta có thể biểu diễn các tham số tiên nghiệm của mô hình bằng biểu đồ phân bố bên dưới.

Hình 5. Biểu đồ phân bố của các tham số tiên nghiệm

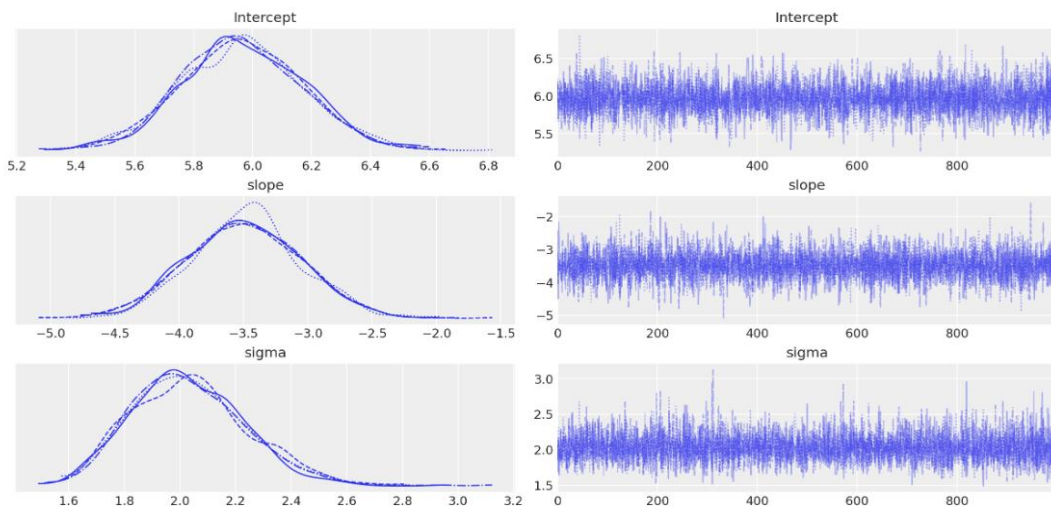


Tham số hậu nghiệm được ước tính bao gồm các giá trị trung bình thực nghiệm và độ lệch chuẩn của từng biến, cộng với sai số chuẩn của giá trị trung bình.

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
sigma	27.09	494.60	0.00	38.99	11.05	7.82	1982.48	1959.93	NaN
Intercept	5.96	0.29	5.39	6.47	0.01	0.00	1935.12	1962.49	NaN
slope	-3.50	0.64	-4.64	-2.27	0.01	0.01	1978.72	1962.81	NaN

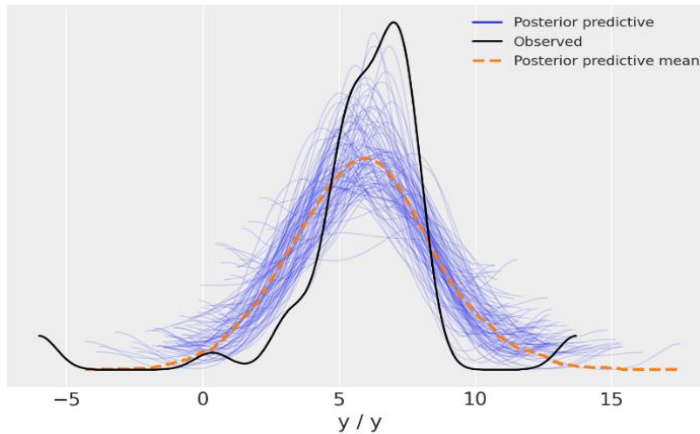
Chúng ta biểu diễn các tham số hậu nghiệm bằng các biểu đồ truy vấn và mật độ sau.

Hình 6. Biểu đồ phân bố của các tham số hậu nghiệm



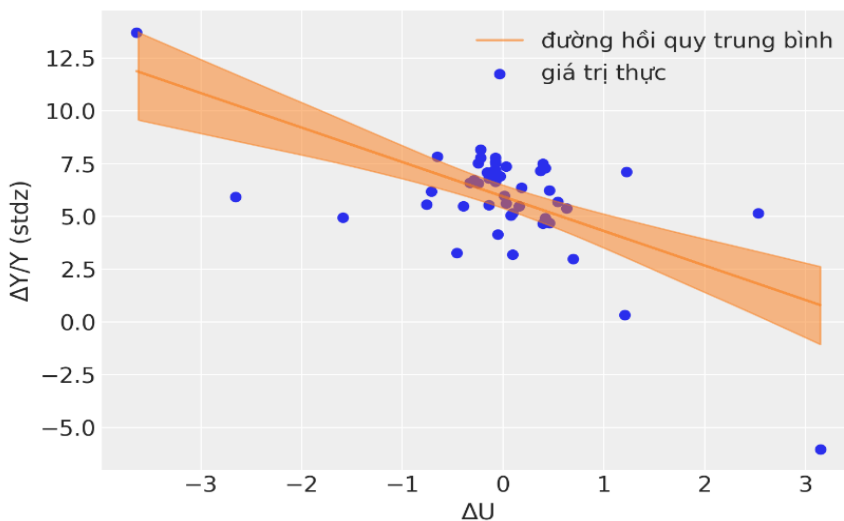
Phân bố hậu nghiệm của dự đoán được trực quan hóa bởi biểu đồ bên dưới:

Hình 7. Phân bố hậu nghiệm của dự đoán thống kê Bayes



Mô phỏng mô hình dự báo hậu nghiệm được trực quan hóa bởi hình dưới:

Hình 8. Mô phỏng mô hình hồi quy Bayes



Hệ số R^2 của mô hình hồi quy Bayes là 0,402. Điều này có nghĩa là mô hình hồi quy Bayes có khả năng giải thích 40,2% biến động của biến phụ thuộc. Mặc dù mô hình Bayes tính đến sự bất định và biến thiên của các tham số nhưng hệ số xác định này bằng với giá trị R^2 của mô hình hồi quy tuyến tính truyền thống. Điều đó cho thấy hai mô hình hồi quy là phù hợp như nhau.

Qua xem xét các mô hình hồi quy thống kê truyền thống, mô hình máy học và mô

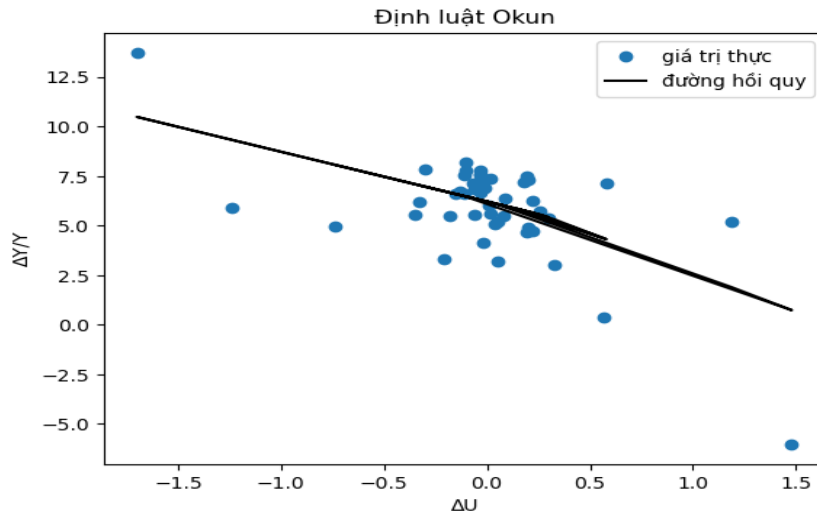
hình Bayes có thể kết luận rằng mô hình hồi quy Keras máy học là lựa chọn tối ưu (hệ số xác định R^2 lớn hơn cả). Định luật Okun trong trường hợp này được viết dưới dạng thức sau:

$$\frac{\Delta Y}{Y} = 0,34 - 0,18\Delta u$$

Định luật Okun cho biết tỷ lệ thất nghiệp theo quý tăng lên 1% khiến tăng trưởng GDP theo quý giảm 0,18%, hay ngược lại, tỷ lệ thất nghiệp giảm 1% khiến tăng trưởng GDP tăng lên 0,18%.

Định luật Okun thể hiện mối tương quan giữa một biến phụ thuộc là tốc độ tăng GDP và một biến độc lập là mức thay đổi tỷ lệ thất nghiệp được thể hiện bởi hình dưới đây.

Hình 9. Mô phỏng định luật Okun (mô hình Keras)



3. Kết luận

Mối quan hệ này được kiểm định bằng cách hồi quy tốc độ tăng trưởng GDP theo mức thay đổi của tỷ lệ thất nghiệp. Tuy nhiên, cũng có những tranh luận rằng thực tế mức giảm sản lượng phần lớn do ảnh hưởng của các yếu tố khác bên cạnh tỷ lệ thất nghiệp. Khi giữ cho các yếu tố khác không đổi, với khung thời gian nghiên cứu từ năm 2011 đến năm 2023, chúng tôi đã ước lượng sản lượng giảm khoảng 0,18% hàng quý tương ứng với mỗi 1% tăng của tỷ lệ thất nghiệp theo quý.

Định luật Okun có ý nghĩa quan trọng trong việc nghiên cứu và dự báo tình hình tiền lương và thất nghiệp trong một nền kinh tế. Nó cũng được sử dụng để đánh giá hiệu quả của chính sách kinh tế và các biện pháp hỗ trợ thị trường lao động. Tuy nhiên, hệ số giữa tăng trưởng kinh tế và tỷ lệ thất nghiệp không nhất quán trong suốt thời gian và có thể thay đổi theo từng quốc gia và thời kỳ

kinh tế. Các kết luận rút ra từ định luật Okun có thể chấp nhận được ở một chừng mực nhất định và là một công cụ phân tích xu hướng ngắn hạn hơn là cho các phân tích dài hạn hay các tính toán đòi hỏi sự chính xác.

Đồng thời chúng tôi muốn khẳng định rằng việc sử dụng các phương pháp thống kê hiện đại đã giúp phát hiện các mối quan hệ và xu hướng trong dữ liệu, suy luận, dự đoán và ra quyết định dựa trên các số liệu thống kê và các mô hình toán học một cách khoa học và chính xác hơn trong nhiều lĩnh vực khác nhau.

Tài liệu tham khảo

1. Định luật Okun - Wikiwand
2. GLM: Linear regression — PyMC 5.10.3 documentation
3. Okun's Law: Economic Growth and Unemployment (investopedia.com)
4. Tensorflow & Keras Tutorial: Linear Regression (learnopencv.com)