

Chương 4-5.

MÔ HÌNH DỰ ĐOÁN CHUỖI THỜI GIAN ĐA BIẾN

PhD. Nguyễn Thị Khánh Tiên

email: tienntk@ut.edu.vn



Nội dung chính trong chương 4:

1. Chuỗi thời gian đơn biến và các đặc trưng quan trọng
2. Nhóm mô hình dự đoán chuỗi thời gian đa biến
 - Mô hình thống kê cổ điển
 - Mô hình Machine Learning
 - Mô hình Deep Learning

Mục tiêu thực hành chương 4:

1. Hiểu rõ bản chất và đặc trưng của chuỗi thời gian đơn biến trong bài toán dự báo.
2. Nắm vững nguyên lý hoạt động của các nhóm mô hình dự đoán chuỗi thời gian
3. Phân tích được điều kiện áp dụng, ưu điểm và hạn chế của từng nhóm mô hình.
4. Biết cách lựa chọn mô hình phù hợp dựa trên đặc tính dữ liệu và mục tiêu bài toán.
5. Chuẩn bị nền tảng lý thuyết cho các chương thực hành dự báo chuỗi thời gian nâng cao.

Khái niệm chuỗi thời gian đa biến (Multivariate Time Series)

Chuỗi thời gian đa biến là tập hợp nhiều chuỗi thời gian quan sát đồng thời theo cùng trục thời gian, có thể tác động lẫn nhau.

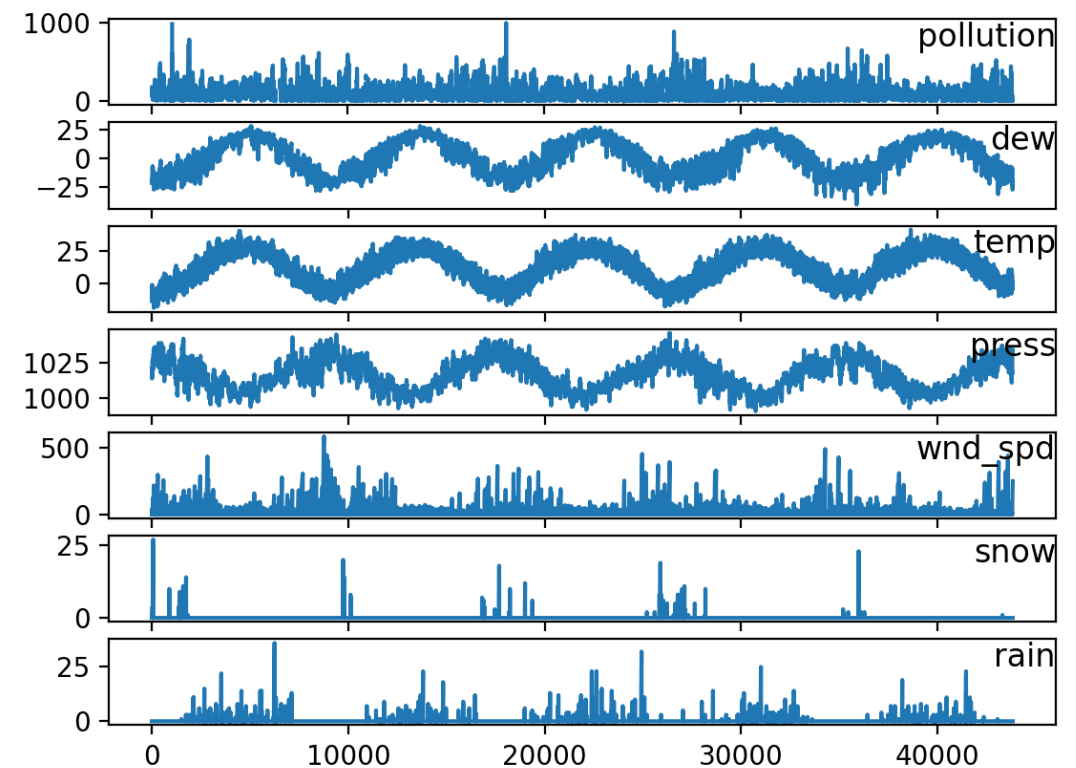
$$\mathbf{X}_t = [x_{1,t}, x_{2,t}, \dots, x_{k,t}]$$

Ví dụ:

- GDP, lạm phát, lãi suất theo quý
- Nhiệt độ, độ ẩm, áp suất
- Lượng bán hàng + giá + khuyến mãi

So sánh chuỗi thời gian đơn biến và đa biến

Tiêu chí	Đơn biến	Đa biến
Số chuỗi	1	≥ 2
Phụ thuộc	Chỉ quá khứ của chính nó	Phụ thuộc chéo
Phân tích	Đơn giản	Phức tạp
Dự báo	Hạn chế	Chính xác hơn nếu biến liên quan



Những điểm chính của dự báo chuỗi thời gian đa biến

Dự báo chuỗi thời gian đa biến không chỉ là dự đoán một chuỗi riêng lẻ, mà là **mô hình hóa động lực của cả một hệ thống gồm nhiều biến tương tác theo thời gian**. Các điểm cốt lõi gồm:

1. Phụ thuộc chéo giữa các biến (Cross-dependence)

Trong chuỗi đa biến, mỗi biến $x_i(t)$ có thể bị ảnh hưởng bởi:

- chính quá khứ của nó,
- quá khứ của các biến khác.

Ví dụ:

$$load_t = f(load_{t-1}, temp_{t-1}, humidity_{t-1}, wind_{t-1})$$

Nếu bỏ qua mối liên hệ này, mô hình sẽ mất đi phần lớn thông tin dự báo.

2. Phụ thuộc theo thời gian (Temporal dependence)

Các biến không chỉ phụ thuộc chéo mà còn có:

- chu kỳ ngày, tuần, mùa
- xu hướng dài hạn
- độ trễ khác nhau giữa các biến

Ví dụ: nhiệt độ ảnh hưởng tới điện năng sau 1–2 giờ, không phải ngay lập tức.

3. Độ trễ khác nhau giữa các biến (Lag heterogeneity)

Mỗi biến có độ trễ tác động khác nhau:

- Nhiệt độ → điện: trễ 1–3 giờ
- Độ ẩm → điện: trễ 6–12 giờ

Do đó, mô hình phải học: $x_i(t) \leftarrow x_j(t - \tau_{ij})$

4. Nhiều và không dừng (Non-stationarity & noise)

Dữ liệu đa biến thường:

- có xu hướng
- có mùa vụ
- bị nhiễu không đồng nhất

Điều này làm các mô hình tuyến tính đơn giản (VAR) dễ thất bại nếu không xử lý đúng.

5. Không gian đặc trưng rất lớn với k biến, L độ trễ \Rightarrow Số đặc trưng là: $k \times L$

Ví dụ: 10 biến \times 48 giờ = 480 đặc trưng \rightarrow dễ quá khớp nếu dữ liệu nhỏ.

6. Bài toán multi-step & vector forecasting

Trong chuỗi đa biến, ta có thể cần dự báo: một biến (load) hoặc cả vector \mathbf{X}_{t+h} , và thường phải dự báo nhiều bước: $[X_{t-24}, \dots, X_t] \rightarrow [X_{t+1}, \dots, X_{t+24}]$
 \Rightarrow lỗi tích lũy rất nhanh.

Mô hình thống kê cổ điển

Mô hình VAR (Vector Autoregression)

- VAR được xây dựng để mở rộng AR sang chuỗi đa biến
- Điều kiện áp dụng: chuỗi dừng (stationary), quan hệ tuyến tính, nếu chuỗi không dừng \rightarrow phải sai phân trước khi huấn luyện
- Cho một chuỗi đa biến:

$$\mathbf{X}_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{k,t} \end{bmatrix}$$

Mô hình VAR(p):

$$\mathbf{X}_t = \mathbf{c} + A_1\mathbf{X}_{t-1} + A_2\mathbf{X}_{t-2} + \dots + A_p\mathbf{X}_{t-p} + \boldsymbol{\varepsilon}_t$$

Trong đó:

- A_i : ma trận hệ số ($k \times k$)
- \mathbf{c} : vector hằng
- $\boldsymbol{\varepsilon}_t$: nhiễu ngẫu nhiên

Ví dụ với 2 biến:

- Load (điện)
- Temperature
- $Load_t = a_1 Load_{t-1} + a_2 Temp_{t-1} + \varepsilon_t$

Nhiệt độ trong quá khứ ảnh hưởng trực tiếp đến điện năng hiện tại

Mô hình thống kê cổ điển

Mô hình VARIMA (Vector ARIMA)

- **VARIMA là phiên bản mở rộng của VAR, tương tự như mối quan hệ giữa AR và ARIMA.**
- VARIMA mạnh hơn VAR về mặt lý thuyết, cho phép mô hình:
 - làm việc với chuỗi đa biến
 - chuỗi có xu hướng, không dừng
 - chuỗi có nhiều phụ thuộc theo thời gian
- **Công thức VARIMA(p,d,q)**

$$\Phi(B) (1 - B)^d \mathbf{X}_t = \Theta(B) \varepsilon_t$$

Trong đó:

- B là toán tử trễ
- d là số lần sai phân
- $\Phi(B)$: thành phần VAR
- $\Theta(B)$: thành phần MA

Những hạn chế của VARIMA khi áp dụng thực tế

- Rất khó ước lượng
- Có khả năng không hội tụ
- Gặp khó khăn khi dữ liệu nhiều biến
- Triển khai phức tạp

Mô hình Machine Learning

Định hình bài toán

Machine Learning tiếp cận chuỗi đa biến bằng cách **chuyển đổi bài toán dự báo thành bài toán hồi quy có giám sát**.
Với vector quan sát:

$$X_t = [x_{1,t}, x_{2,t}, \dots, x_{k,t}]$$

và cửa sổ trễ độ dài L , đầu vào tại thời điểm t được xây dựng từ:

$$[X_{t-1}, X_{t-2}, \dots, X_{t-L}]$$

Mục tiêu dự đoán:

$$y_t = x_{j,t+1}$$

Feature Engineering cho chuỗi đa biến

Hiệu quả của mô hình ML phụ thuộc mạnh vào cách biểu diễn dữ liệu:

- Lag features: giá trị quá khứ của tất cả các biến
- Rolling statistics: trung bình, độ lệch chuẩn, min, max theo cửa sổ thời gian
- Cross-variable features: tỷ lệ, hiệu, hoặc tương tác giữa các biến (ví dụ: load / temperature)

Các mô hình Machine Learning tiêu biểu

- Linear Regression
- Decision Tree, Random Forest
- Gradient Boosting (XGBoost, LightGBM, CatBoost)

Vai trò

- Machine Learning đặc biệt phù hợp khi: số biến lớn; quan hệ giữa các biến phức tạp; chuỗi có mùa vụ, xu hướng, nhiễu phi tuyến
- => Do đó, các mô hình như **XGBoost** đã trở thành tiêu chuẩn thực tế trong các hệ thống dự báo chuỗi thời gian đa biến **dạng tabular**.

Mô hình Deep Learning

Biểu diễn chuỗi đa biến cho Deep Learning

Một chuỗi đa biến được biểu diễn dưới dạng tensor:

$$X \in \mathbb{R}^{T \times K}$$

với:

- T : số bước thời gian
- K : số biến

Với cửa sổ trễ độ dài L , đầu vào của mô hình có dạng:

$$X_t = [X_{t-L+1}, \dots, X_t] \in \mathbb{R}^{L \times K}$$

Mục tiêu là dự đoán:

$$X_{t+h}$$

hoặc một tập các bước tương lai.

Các mô hình Machine Learning tiêu biểu

- Recurrent Neural Networks (RNN, LSTM, GRU)
- Temporal Convolutional Networks (TCN)
- Transformer-based Models

Vai trò

Trong thực tế, các mô hình như **LSTM**, **TCN** và **Transformer** đã thay thế VAR và ARIMA trong nhiều hệ thống lớn vì:

- Không cần giả định dừng
- Tự học biểu diễn
- Mở rộng tốt cho dữ liệu lớn và đa biến

Các mô hình học sâu đã trở thành công cụ chủ đạo để mô hình hoá hệ thống dự đoán nhờ khả năng học quan hệ phi tuyến, phụ thuộc dài hạn và cấu trúc đa chiều

So sánh Deep Learning và XGBoost cho chuỗi thời gian đa biến

Tiêu chí	XGBoost (Machine Learning)	Deep Learning (LSTM / TCN / Transformer)
Biểu diễn dữ liệu	Dạng bảng (tabular) với lag, rolling, cross-features	Tensor 3D ((samples, time, variables))
Feature engineering	Bắt buộc, rất quan trọng	Gần như không cần
Học phụ thuộc theo thời gian	Gián tiếp qua lag	Trực tiếp trong kiến trúc
Phụ thuộc dài hạn	Kém khi lag dài	Rất tốt
Quan hệ giữa các biến	Qua chia nhánh cây	Qua không gian ẩn nhiều tầng
Chuỗi dài	Không hiệu quả	Rất hiệu quả
Nhiều biến (k lớn)	Bắt đầu khó khi $k > 50$	Hoạt động tốt với hàng trăm biến
Multi-step forecasting	Khó, phải lặp	Tự nhiên, thiết kế sẵn
Dữ liệu nhỏ	Rất tốt	Dễ overfit
Dữ liệu lớn	Trung bình	Rất tốt
Độ chính xác	Cao cho chuỗi tabular	Cao cho chuỗi phức tạp
Overfitting	Ít hơn	Dễ hơn
Giải thích mô hình	Feature importance, SHAP	Rất khó
Tốc độ huấn luyện	Nhanh	Chậm
Tài nguyên tính toán	CPU là đủ	Thường cần GPU
Triển khai thực tế	Dễ	Phức tạp
Ứng dụng điển hình	Bán hàng, điện năng, kinh tế	Giao thông, IoT, tài chính, sensor

So sánh các kiến trúc Deep Learning cho chuỗi thời gian đa biến

Tiêu chí	LSTM / GRU	TCN (Temporal CNN)
Nguyên lý	Recurrent + Memory cell	Causal & dilated convolution
Biểu diễn dữ liệu	Chuỗi tuần tự	Chuỗi như tín hiệu
Phụ thuộc dài hạn	Tốt	Tốt
Phụ thuộc ngắn hạn	Tốt	Rất tốt
Song song hóa	Kém	Tốt
Tốc độ huấn luyện	Chậm	Nhanh
Ổn định gradient	Trung bình	Tốt
Multi-step forecasting	Tốt	Tốt
Nhiều biến (k lớn)	Trung bình	Tốt
Chuỗi rất dài	Kém	Tốt
Dữ liệu cần	Trung bình	Trung bình
Overfitting	Dễ	Trung bình
Giải thích	Rất khó	Rất khó
Ứng dụng tiêu biểu	Điện năng, bán hàng	Sensor, IoT

Reference

1. <https://machinelearningmastery.com/multivariate-time-series-forecasting-lstms-keras/>
2. <https://www.geeksforgeeks.org/data-analysis/time-series-forecasting-using-pytorch/>
3. <https://www.kaggle.com/code/davidandressanchez/xgboost-for-multivariate-ts>
4. <https://www.kaggle.com/code/sajikim/time-series-forecasting-methods-example-python>
5. <https://machinelearningmastery.com/how-to-develop-machine-learning-models-for-multivariate-multi-step-air-pollution-time-series-forecasting/>