HỌ MÔ HÌNH TUYẾN TÍNH MỞ RỘNG VỚI MÔ HÌNH NHIỄU TUYẾN TÍNH

M. Sc. VŨ VĂN TUẤN
Vụ KHKT

1. ĐẶT VẤN ĐỀ

Kể từ khi Sherman đề xuất khái niệm về đường đơn vị, phương pháp phân tích hệ thống tuyến tính đã đóng một vai trò quan trọng trong lịch sử phát triển của thủy văn học, đặc biệt là trong kỹ thuật mô hình hóa mưa dòng chảy và trong diễn toán lũ. Giả thuyết về thủy đồ đơn vị dựa trên cơ sở giả định về sự phù hợp của nguyên lý chồng chất (superposition) trong một hệ thống tuyến tính bất biến theo thời gian, cho phép biểu diễn mối liên hệ giữa lượng mưa hữu hiệu x(t) và lượng dòng chảy y (t) qua tích phân chập sau :

$$y(t) = \int_0^{\infty} (\tau) . h(t-\tau) . d\tau \qquad (1)$$

trong đó : τ là biến của tích phân ; h (t) là hàm phản ứng mạch động (tung độ của đường đơn vị tức thời) tại thời điểm t.

Tuy nhiên, về bản chất, các quá trình mưa – dòng chảy hay diễn toán lũ đều không phải là tuyến tính. Chúng ta đã biết, diễn toán dòng chảy trong kênh hở tuân theo hệ phương trình Saint Venant mang tính chất phi tuyến. Cũng như vậy, trong quá trình mưa — dòng chảy, lượng mưa hữu hiệu phụ thuộc rất nhiều vào trạng thái của lưu vực khi xảy ra mưa (như trạng thái ẩm và khả năng thấm của đất), điều kiện này lại phụ thuộc vào chế độ mưa trước đó (như lượng và quá trình diễn biến mưa, thời gian gián đoạn giữa hai trận mưa..) cũng như khả năng bốc hơi của lưu vực giữa hai trận mưa... mà những điều kiện này lại biến đổi mạnh theo mùa. Rõ ràng là những điều đó khó có thể được đáp ứng trong một hệ thống tuyến tính mang tính chất tỷ lệ và bất biến theo thời gian.

Các nhà thủy văn đã sớm nhận thấy điều đó và đã có gắng tìm các hướng giải quyết mới. Năm 1963, Amorocho [1] đã gợi ý thay thế (1) bởi chuỗi hàm :

$$y(t) = \int_0^t x(\tau)h(t,\tau)d\tau + \int_0^t x(\tau_2) \int_0^t x(\tau_1)h(t, \tau_1, \tau_2)d\tau_1 d\tau_2 + ... \quad (2)$$

trong đó số hạng tích phân thứ nhất biểu diễn hoạt động của hệ thống « xấp xỉ » tuyến tính, còn các số hạng tích phân sau là những tích phân chập có bậc cao hơn, đặc trưng cho tính chất phi tuyến của hệ.

Tuy nhiên, mặc dù các mô hình phi tuyến có tạo ra được những hướng phát triển mới song vẫn không thay thế được những mô hình tuyến tính. Xuất phát từ những giả thiết đơn giản và có tính thực dụng cao, cũng như sự đảm bảo về độ chính xác trong tính toán và dự báo, các mô hình tuyến tính vẫn đang được tiếp tục phát triển [3].

2. MÔ HÌNH TUYẾN TÍNH ĐƠN VỚI MỘT INPUT

Với những hàm liên tục, quan hệ input – ouput của một hệ thống tuyến tính tập trung, bất biến theo thời gian có thể được biểu diễn theo (1). Khi hàm input được biểu diễn như là một chuỗi các xung hoặc các giá trị trung bình trong thời đoạn T thì ta có tổng chập tương ứng sau:

$$y_i = x_i h_1 + x_{i-1} h_2 + ... + x_{i-m+1} h_m + e_i$$

hay:

$$y_i = \sum_{j=1}^{m} x_{i-j+1} h_j + e_i \quad (3)$$

trong đó: h_j biểu diễn tung độ thứ j của phản ứng mạch động; m là độ dài bộ nhớ của hệ thống (nghĩa là tác động của input x chỉ kéo dài qua m thời đoạn có độ dài mỗi thời đoạn là T); e_i là số hạng sai số trong quan hệ.

Phương trình (3) có thể được diễn giải chi tiết hơn theo các tung độ y_i với n phương trình đại số tuyến tính:

$$\begin{aligned}
y_1 &= x_1 h_1 && + e_1 \\
y_2 &= x_2 h_1 + x_1 h_2 && + e_2 \\
y_3 &= x_3 h_1 + x_2 h_2 + x_1 h_3 && + e_3 \\
&\cdots\cdots\cdots \\
y_m &= x_m h_1 + x_{m-1} h_2 + ... + x_1 h_m && + e_m \\
y_{m+1} &= x_{m+1} h_1 + x_m h_2 + ... + x_2 h_m && + e_{m+1} \\
&\cdots\cdots \\
y_n &= x_n h_1 + x_{n-1} h_2 + ... + x_{n-m+1} h_m && + e_n
\end{aligned}$$

Nếu biểu diễn dưới dạng tích ma trận – véc tơ, ta có:

$$
\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ y_m \\ y_{m+1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}
=
\begin{bmatrix} x_1 & 0 & 0 & \ldots & 0 \\ x_2 & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_m & x_{m-1} & & \ldots & x_1 \\ x_{m+1} & x_m & & \ldots & x_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_n & x_{n-1} & & \ldots & x_{n-m+1} \end{bmatrix}
\begin{bmatrix} h_1 \\ h_2 \\ \cdot \\ \cdot \\ h_m \end{bmatrix}
+
\begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_m \\ \cdot \\ \cdot \\ e_n \end{bmatrix}
\qquad (4)
$$

trong đó: m là độ dài bộ nhớ của hệ thống, n, là số các quan trắc của biến y. Cũng có thể biểu diễn (4) dưới dạng rút gọn:

$$Y = X.H + E \qquad (5)$$

trong đó: Y là véctơ cột (n, 1) của chuỗi output,

X là ma trận (n, m) của chuỗi input,

H là véctơ cột (m, 1) của các tung độ hàm phản ứng mạch động

E là véctơ cột (n, 1) của các sai số.

Nghiệm của (5) đã được giải trên cơ sở cực tiểu của tổng bình phương các sai số:

$$\min \left\{ \sum_{i=1}^{n} e_i^2 \right\}$$

dẫn đến nghiệm:

$$\hat{H} = [X^T X]^{-1} . X^T Y \qquad (6)$$

với $\hat{H}$ là nghiệm theo nghĩa bình phương tối thiểu của vectơ H, X^T là ma trận chuyển vị của ma trận X và $[X^T X]^{-1}$ là ma trận nghịch đảo của ma trận tích $[X^T X]$

3. MÔ HÌNH TUYẾN TÍNH ĐƠN VỚI NHIỀU INPUT

Đối với những hệ thống sông lớn, thường ta cần tiến hành tính toán – dự báo trên cơ sở có nhiều input (ngoài mưa ra còn có một số dòng chảy nhập lưu–mà đôi khi so với chúng thì input « mưa khu giữa » chỉ chiếm một tỷ lệ nhỏ). trong trường hợp này, chúng ta có thể sử dụng mô hình tuyến tính đơn với nhiều input.

Một hệ thống biểu diễn mối quan hệ giữa một output với nhiều input được gọi là tuyến tính nếu quan hệ của mỗi input đối với thành phần tương ứng của nó trong output là tuyến tính và nếu nguyên lý chồng chất áp dụng được cho các tổ hợp của các thành phần trong output. Như vậy, đối với một hệ thống nhiều input – một output, có thể biểu diễn dưới dạng:

$$y_t = \sum_{k=1}^{L} \int_0^{\infty} h^{(k)}(\tau).x^{(k)}(t-\tau)d\tau \tag{7}$$

trong đó: L là tổng số các hàm input, $x^{(k)}$ là hàm input thứ k, $h^{(k)}$ là hàm phản ứng mạch động đơn vị tương ứng đối với input thứ k

Rõ ràng khi L = 1, đẳng thức (7) trở về dạng (1). Trong trường hợp biến rời rạc, chúng ta cũng có tổng chập tương ứng:

$$y_i = \sum_{k=1}^{L} \sum_{j=1}^{m(k)} x^{(k)}_{i-j+1} h^{(k)}_j \tag{8}$$

trong đó m(j) là độ dài bộ nhớ của hệ thống tương ứng với chuỗi input thứ j.

4. Mô hình nhiễu tuyến tính

Như đã nêu trên, đối với các mô hình tính toán – dự báo thời đoạn ngắn (như trận lũ), phải tính đến ảnh hưởng của sự thay đổi trong điều kiện ẩm của lưu vực. Còn đối với các mô hình tính toán – dự báo thời đoạn dài (quá trình năm), phải tính đến sự biến đổi theo mùa của khả năng bốc hơi. Việc đưa vào những thành phần phi tuyến như chuỗi hàm (2) nêu trên là một hướng. Tuy nhiên, trong họ các mô hình tuyến tính cũng có thể xử lý được tính biến đổi theo thời gian nhờ những khái niệm tuyến tính mở rộng như J.E.Nash đã đề xuất thông qua mô hình nhiễu tuyến tính. Mô hình này dựa trên hai giả định sau đây [2,3]:

a) Nếu trong một năm nào đó, mỗi một hàm input nhận giá trị trong ngày thứ d bằng giá trị kỳ vọng của nó thì hàm output cũng nhận giá trị kỳ vọng của nó tại ngày thứ d, nghĩa là $i_d \rightarrow q_d$.

b) Nhiễu (hay độ lệch khỏi giá trị kỳ vọng) của input có quan hệ tuyến tính với nhiễu của output, nghĩa là $(i - i_d) \rightarrow (q - q_d)$ trong đó: i_d, q_d là các giá trị kỳ vọng của input và output tại ngày thứ d; i, q là các giá trị thực của input và output tại ngày thứ d; $(i - i_d)$, $(q - q_d)$ là các nhiễu của input và output tại ngày thứ d.

Như vậy, để tiến hành tính toán – dự báo theo mô hình nhiễu tuyến tính, trước hết phải xác định được mô hình mùa của chuỗi các input và output. Mô hình mùa dự báo cho ngày thứ d của năm (d biến đổi từ 1 đến 365) theo giá trị trung bình của thời kỳ định chuẩn:

$$\hat{y}_d = \frac{1}{L} \sum_{r=1}^{L} y_{d,r} \tag{9}$$

trong đó: $y_{d,r}$ là dòng chảy quan trắc trong ngày thứ d của năm thứ r, còn L là số năm trong thời kỳ định chuẩn.

Với L lớn, các giá trị tính toán của y_d dao động mạnh, do đó cần được làm trơn. Có thể sử dụng chuỗi Fourier cho mục đích này:

$$y_d = a_0 + \sum_{j=1}^{p} A_j Cos(\frac{2\pi_j d}{n})\ B_j Sin(\frac{2\pi_j d}{n}) \tag{10}$$

trong đó: a_0 là giá trị trung bình của y_d; A_j; B_j là các hệ số Fourier; $_j$ là bậc của các điều hòa; p là số điều hòa cực đại; n là bằng 365 cho chuỗi số liệu ngày.

Trên đây là các nghiệm phi thông số của các mô hình tuyến tính đơn và mô hình nhiễu tuyến tính (có thể áp dụng cho cả hai trường hợp có một hay nhiều input).

Tài liệu tham khảo

1. Vũ Văn Tuấn. Liear Modelling Techniques and Applications. International Workshop on River Flow Forecasting, DCG, Ireland (1989).

2. Amorocho. J. Measures of the Linearity of Hydrologic Systems. J.Geological Research, No.68 (1963).

3. Nash, J.E. and Barsi, B.I. A Hybrid Model for Flow Forecasting on Large Catchments.J.Hydrol. No.65 (1983).