

# Các mô hình dự báo áp dụng trong phần mềm

## 1. Mô hình Holt-Winters

Một chuỗi thời gian bất kỳ (dữ liệu của một biến kinh tế là một chuỗi thời gian) thường có 4 thành phần quan trọng: xu thế, chu kỳ, mùa vụ và bất quy tắc, trong đó hai thành phần xu thế và mùa vụ là quan trọng nhất.

- Với chuỗi thời gian: không có xu thế và mùa vụ, người ta có phương pháp (hay mô hình) làm trơn hàm mũ đơn để dự báo;
- Nếu có tính xu thế: đề xuất mô hình làm trơn hàm mũ kép;
- Nếu có cả tính xu thế và tính mùa vụ: đề xuất mô hình làm trơn hàm mũ bậc 3, còn gọi là mô hình Holt-Winter. Nếu xu thế và mùa vụ kết hợp với nhau theo kiểu nhân: ta có mô hình Holt\_Winter mùa vụ nhân, kết hợp theo kiểu cộng: ta có mô hình Holt-Winter mùa vụ cộng;
- Ngoài ra người ta còn sử dụng cách tiếp cận Holt-Winter cho chuỗi thời gian có tính xu thế, nhưng không có tính mùa vụ: mô hình Holt-Winter không mùa vụ. Mô hình Holt-Winter không mùa vụ có chất lượng dự báo tốt hơn Mô hình làm trơn hàm mũ kép.
- Phương pháp làm trơn hàm mũ có 5 loại mô hình, nhưng thực tế Mô hình kép nói chung có chất lượng kém hơn Mô hình Holt-Winter không mùa vụ nên: thực tế người ta chỉ cần 4 mô hình là đủ.

Thuật toán

Giả sử: giá trị thực tế thu thập được tại thời điểm  $t$  của một biến kinh tế vĩ mô nào đó, ta cần dự báo cho 2 kỳ tiếp theo của biến kinh tế này. Thuật toán xây dựng phần mềm được dựa trên phương pháp làm trơn hàm mũ chuỗi thời gian.

Thực hiện phương pháp làm trơn hàm mũ chuỗi thời gian để nhận được chuỗi dự báo (hay chuỗi làm trơn phù hợp nhất). Mô hình dự báo một chuỗi thời gian với chuỗi bất kỳ sẽ chỉ rơi vào một trong 5 (hay 4) dạng dưới đây. Mô hình nào cho RMSE nhỏ nhất là mô hình được chọn.

### 1.1. Làm trơn hàm mũ bậc 1

Công thức của phương pháp làm trơn hàm mũ dạng đơn giản:

$$\bar{y}_t = \alpha * y_t + (1 - \alpha) * \bar{y}_{t-1} \quad 0 \leq \alpha \leq 1$$

hay viết cách khác:  $\bar{y}_t = \bar{y}_{t-1} + \alpha * (y_t - \bar{y}_{t-1})$ , trong đó:

$\bar{y}_t$ : giá trị ước lượng tại thời điểm  $t$

$y_t$ : giá trị thực tế thu thập được tại thời điểm  $t$ ;

$\bar{y}_{t-1}$ : giá trị ước lượng tại thời điểm trước đó

$\alpha$ : hằng số làm trơn.

-  $F_{m+k}$  là giá trị dự báo của  $y_{m+k}$  được thực hiện tại thời điểm  $m$ .  $F_{m+k} = \bar{y}_m$ .

### Thuật toán lập trình:

a. Khởi đầu tại  $t=1$ ,  $\bar{y}_1 = y_1$ ;

b. Tính lặp:  $\bar{y}_t = \alpha * y_t + (1-\alpha) * \bar{y}_{t-1}$ ,  $t=2, 3, \dots, M$ ;

c. Dự báo:  $F_{M+k} = \bar{y}_M$ . ( $k = 1, 2, 3, \dots$ ).

- Hệ số  $\alpha$ : là nghiệm của bài toán tối ưu được thực hiện bằng tính toán lưới (grid calculate) dưới đây:

**Giải bài toán:** Tìm  $\alpha$  sao cho:

$$\text{MSE} = \frac{1}{m} \sum_{t=1}^m (y_{t+k} - \bar{y}_{t+k})^2 \rightarrow \text{Min}$$

$$\bar{y}_t = \alpha * y_t + (1-\alpha) * \bar{y}_{t-1} \quad 0 \leq \alpha \leq 1$$

hay viết cách khác:  $\bar{y}_t = \bar{y}_{t-1} + \alpha * (y_t - \bar{y}_{t-1})$ , trong đó:

$\bar{y}_t$ : giá trị ước lượng tại thời điểm  $t$

$y_t$ : giá trị thực tế thu thập được tại thời điểm  $t$ ;

$\bar{y}_{t-1}$ : giá trị ước lượng tại thời điểm trước đó

$\alpha$ : tham số làm trơn.

## 1.2. Làm trơn hàm mũ bậc 2 theo cách tiếp cận của Holt-Winters

Thuật toán lập trình

a. Khởi đầu tại  $t = 2$ ,  $\bar{y}_2 = y_2$ ;  $T_2 = y_2 - y_1$  hoặc  $T_2 = (y_4 - y_1)/3$  hoặc  $T_2 = (y_n - y_1)/(n-1)$ .

b. Lặp:  $\bar{y}_t = \alpha * y_t + (1-\alpha) * (\bar{y}_{t-1} + T_{t-1})$   $0 \leq \alpha \leq 1$

$$T_t = \beta * (\bar{y}_t - \bar{y}_{t-1}) - (1-\beta) * T_{t-1} \quad 0 \leq \beta \leq 1, t = 3, 4, \dots, m.$$

c. Dự báo:  $F_{m+k} = \bar{y}_m + kT_m$

Ở đây:

$F_{m+k}$ : giá trị dự báo trước k bước mắt đầu từ thời điểm  $m$ ,  $k = 1, 2, \dots$

$y_t$ : giá trị thực tế thu thập được tại thời điểm  $t$ ;

$\bar{y}_t$ : giá trị được ước lượng (hay được làm tròn) của chuỗi dữ liệu ban đầu tại thời điểm  $t$

$T_t$ : giá trị làm tròn xu thế tại thời điểm  $t$

$\alpha, \beta$ : tham số làm tròn là nghiệm của bài toán tối ưu sau:

**Giải bài toán:** Tìm  $a, b$  sao cho:

$$\text{MSE} = \frac{1}{m} \sum_{t=1}^m (y_{t+k} - \bar{y}_{t+k})^2 \rightarrow \text{Min}$$

$$\bar{y}_t = \alpha * y_t + (1 - \alpha) * (\bar{y}_{t-1} + T_{t-1}) \quad 0 \leq \alpha \leq 1$$

$$T_t = \beta * (\bar{y}_t - \bar{y}_{t-1}) - (1 - \beta) * T_{t-1}$$

$$0 \leq \beta \leq 1, t = 3, 4, \dots, m.$$

trong đó:

$\bar{y}_t$ : giá trị ước lượng tại thời điểm  $t$

$y_t$ : giá trị thực tế thu thập được tại thời điểm  $t$ ;

$\bar{y}_{t-1}$ : giá trị ước lượng tại thời điểm trước đó ;

### 1.3. Làm tròn Holt-Winters mùa vụ nhân

Thuật toán lập trình: mùa vụ với độ dài  $s$ , được tìm bằng phân tích hàm tự tương quan và tự tương quan riêng phần;

a. **Khởi tạo:** tại  $t = s$ ;

$$\bar{y}_s = \frac{1}{s} \sum_{t=1}^s y_t$$

$$T_s = 0;$$

$$S_j = \frac{y_j}{\frac{1}{s} \sum_{t=1}^s y_t}, j = 1, 2, \dots, s$$

b. **Lập**

$$\bar{y}_t = \alpha \frac{y_t}{S_{t-s}} + (1-\alpha) * (\bar{y}_{t-1} + T_{t-1}), \text{ với } 0 \leq \alpha \leq 1$$

$$T_t = \beta * (\bar{y}_t - \bar{y}_{t-1}) + (1-\beta) * T_{t-1}, \text{ với } 0 \leq \beta \leq 1.$$

$$S_t = \gamma * \left(\frac{y_t}{\bar{y}_t}\right) + (1-\gamma) * S_{t-s} \text{ với } 0 \leq \gamma \leq 1.$$

$$t = s+1, \dots, m.$$

### c. Dự báo

$$F_{m+k} = (\bar{y}_m + k * T_m) S_{m+k-s}, k = 1, 2, \dots, s;$$

$$\text{và } F_{m+k} = (\bar{y}_m + k * T_m) S_{m+k-2s}, k = s+1, \dots, 2s.$$

Ở đây: s là độ dài của mùa vụ;

$T_t$  là giá trị làm trơn xu thế;  $S_t$  là giá trị làm trơn mùa vụ;  $F_{m+k}$  là ký hiệu giá trị dự báo trước k bước ngoài mẫu tại thời điểm  $m$ ; các ký hiệu còn lại như đã được giải thích.

- Hệ số  $\alpha, \beta, \gamma$ : là nghiệm của của bài toán tối ưu sau:

**Giải bài toán:** Tìm a, b, g sao cho:

$$\text{MSE} = \frac{1}{m} \sum_{t=1}^m (y_{t+k} - \bar{y}_{t+k})^2 \rightarrow \text{Min}$$

$$\bar{y}_t = \alpha \frac{y_t}{S_{t-s}} + (1-\alpha) * (\bar{y}_{t-1} + T_{t-1}), \text{ với } 0 \leq \alpha \leq 1$$

$$T_t = \beta * (\bar{y}_t - \bar{y}_{t-1}) + (1-\beta) * T_{t-1}, \text{ với } 0 \leq \beta \leq 1.$$

$$S_t = \gamma * \left(\frac{y_t}{\bar{y}_t}\right) + (1-\gamma) * S_{t-s} \text{ với } 0 \leq \gamma \leq 1.$$

trong đó:

$\bar{y}_t$ : giá trị ước lượng tại thời điểm  $t$

$y_t$ : giá trị thực tế thu thập được tại thời điểm  $t$ ;

$\bar{y}_{t-1}$ : giá trị ước lượng tại thời điểm trước đó.

## 1.4. Làm trơn Holt-Winters mùa vụ cộng

Thuật toán lập trình: mùa vụ với độ dài s:

a. **Khởi tạo:** tại  $t = s$ ;

$$\bar{y}_s = \frac{1}{s} \sum_{t=1}^s y_t$$

$$T_s = 0;$$

$$S_j = \frac{y_j}{\frac{1}{s} \sum_{t=1}^s y_t}, j = 1, 2, \dots, s.$$

b. **Lặp**

$$\bar{y}_t = \alpha(y_t - S_{t-s}) + (1-\alpha) * (\bar{y}_{t-1} + T_{t-1}), \text{ với } 0 \leq \alpha \leq 1$$

$$T_t = \beta * (\bar{y}_t - \bar{y}_{t-1}) + (1-\beta) * T_{t-1}, \text{ với } 0 \leq \beta \leq 1.$$

$$S_t = \gamma * (y_t - \bar{y}_t) + (1-\gamma) * S_{t-s} \text{ với } 0 \leq \gamma \leq 1.$$

$$t = s+1, \dots, m.$$

c. **Dự báo**

$$F_{m+k} = \bar{y}_m + k * T_m + S_{m+k-s}, k = 1, 2, \dots, s;$$

$$\text{và } F_{m+k} = \bar{y}_m + k * T_m + S_{m+k-2s}, k = s+1, \dots, 2s.$$

Ở đây:  $s$  là độ dài của mùa vụ;

$T_t$  là giá trị làm trơn xu thế;  $S_t$  là giá trị làm trơn mùa vụ;  $F_{m+k}$  là ký hiệu giá trị dự báo trước  $k$  bước ngoài mẫu tại thời điểm  $m$ ; các ký hiệu còn lại như đã được giải thích.

a, b, g là nghiệm của bài toán tối ưu sau:

**Giải bài toán:** Tìm a, b, g sao cho:

$$\text{MSE} = \frac{1}{m} \sum_{t=1}^m (y_{t+k} - \bar{y}_{t+k})^2 \rightarrow \text{Min}$$

$$\bar{y}_t = \alpha(y_t - S_{t-s}) + (1-\alpha) * (\bar{y}_{t-1} + T_{t-1}), \text{ với } 0 \leq \alpha \leq 1$$

$$T_t = \beta * (\bar{y}_t - \bar{y}_{t-1}) + (1-\beta) * T_{t-1}, \text{ với } 0 \leq \beta \leq 1.$$

$$S_t = \gamma * (y_t - \bar{y}_t) + (1-\gamma) * S_{t-s} \text{ với } 0 \leq \gamma \leq 1.$$

$$t = s+1, \dots, m.$$

trong đó:

$\bar{y}_t$ : giá trị ước lượng tại thời điểm  $t$

$y_t$ : giá trị thực tế thu thập được tại thời điểm  $t$ ;

$\bar{y}_{t-1}$ : giá trị ước lượng tại thời điểm trước đó ;

$\alpha$ ,  $b$ ,  $g$ : là các tham số làm tron.

## 2. Mô hình Tự hồi quy bậc p

Cho chuỗi  $Y_t$  bất kỳ, mô hình dự báo theo phương pháp tự hồi quy bậc p được xây dựng bằng phương pháp ước lượng bình phương tối thiểu OLS phương trình sau đây:

$$\ln \bar{Y}_t = A_1 \ln(\bar{Y}_t(-1)) + \dots + A_p \ln(\bar{Y}_t(-p)) + C + B_1 t + B_2 t^2 + B_3 t^3 + B_4 t^4 + B_5 t^5 + \frac{U(-1) + \dots + U(-p)}{p} + d_1 * \text{Dummy}_1 + \dots + d_h * \text{Dummy}_h + d_{\text{Crisis1997}} * \text{Dummy}_{\text{Crisis1997}} + d_{\text{Crisis2007}} * \text{Dummy}_{\text{Crisis2007}} + e_t$$

để nhận được các tham số phù hợp nhất.

Giống như mô hình dự báo được xây dựng theo phương pháp nhip tăng, điều quan trọng nhất là phải xử lý phần dư  $e_t$  sao cho nó thỏa mãn 5 điều kiện (không có hiện tượng đa cộng tuyến, phần dư có phân phối chuẩn, có kỳ vọng bằng 0, có phương sai không đổi, phần dư không tự tương quan đến bậc k bất kỳ và không tương quan với biến độc lập).

Để xử lý được phần dư của mô hình có tính chất như vậy cần có sự tham gia của con người. Chuyên gia người cũng có thể sử dụng các phần mềm công cụ như EVIEW, STATA, SAS,... để ước lượng mô hình, xử lý phần dư để cuối cùng nhận được phương trình ước lượng tốt nhất. Dựa vào phương trình ước lượng này để lập chương trình dự báo. Như vậy phần mềm dự báo chỉ tiêu kinh tế vĩ mô cũng cần là phần mềm HYBRID. Tức là cần có sự tham gia của con người

## 3. Mô hình Nhip tăng

Cho chuỗi  $Y_t$  bất kỳ, mô hình dự báo theo phương pháp nhip tăng được xây dựng bằng phương trình sau đây:

$$\ln \bar{Y}_t = B_0 + B_1 t + B_2 t^2 + B_3 t^3 + B_4 t^4 + B_5 t^5 + A_1 * \text{AR}(1) + \dots + A_p * \text{AR}(p) + \frac{U(-1) + \dots + U(-p)}{p} + d_1 * \text{Dummy}_1 + \dots + d_h * \text{Dummy}_h + e_t$$

$$\begin{aligned}
&= \mathbf{B}_0 + \mathbf{B}_1 t + \mathbf{B}_2 t^2 + \mathbf{B}_3 t^3 + \mathbf{B}_4 t^4 + \mathbf{B}_5 t^5 + \mathbf{A}_1 \ln(\bar{Y}_t(-1)) + \dots + \mathbf{A}_p * (\ln(\bar{Y}_t(-1))) \\
&+ \dots + \ln(\bar{Y}_t(-p))) + \frac{U(-1) + \dots + U(-p)}{p} + \mathbf{d}_1 * \mathbf{Dummy}_1 + \dots + \mathbf{d}_h * \mathbf{Dummy}_h + \\
&\mathbf{d}_{Crisis1997} * \mathbf{Dummy}_{Crisis1997} + \mathbf{d}_{Crisis2007} * \mathbf{Dummy}_{Crisis2007} + \mathbf{e}_t
\end{aligned}$$

#### 4. Thuật toán kết hợp dự báo

Giả sử  $y_{t+k,t}^a$ ,  $y_{t+k,t}^b$  và  $y_{t+k,t}^c$  là kết quả dự báo của  $y_{t+k}$  tương ứng với ba mô hình dự báo theo phương pháp làm trơn của tự hồi quy bậc p, nhíp tăng và Holt-Winters.

Kết hợp dự báo được xác định trên cơ sở ước lượng phương trình hồi qui sau:

$$\begin{aligned}
y_{t+k,t} = & \beta_0 + \beta_a y_{t+k,t}^a + \beta_b y_{t+k,t}^b + \beta_c y_{t+k,t}^c + \beta_{aa} (y_{t+k,t}^a)^2 + \beta_{bb} (y_{t+k,t}^b)^2 + \beta_{cc} (y_{t+k,t}^c)^2 + \\
& \beta_{ab} y_{t+k,t}^a y_{t+k,t}^b + \beta_{ac} y_{t+k,t}^a y_{t+k,t}^c + \beta_{bc} y_{t+k,t}^b y_{t+k,t}^c + \varepsilon_{t+k,t}
\end{aligned}$$

Ở đây t là biến thời gian. Từ phương trình này ta nhận được kết quả dự báo cuối cùng cho 2 kỳ dữ liệu tiếp theo.