

Một cách tìm hàm xu thế nhờ tích vô hướng

VÕ MINH VINH

1. Mở đầu

Trong phân tích kinh tế, ta thường phải tìm hàm xu thế một đa thức bậc 1, bậc 2 hay bậc 3. Hàm xu thế phải phản ánh gần đúng nhất các số liệu thống kê hàng tháng, hàng quý hay hàng năm. Hàm xu thế càng gần đúng bao nhiêu càng giúp cho việc phân tích, dự báo có hiệu quả bấy nhiêu. Trong bài báo này, chúng tôi trình bày một tìm hàm xu thế dựa trên một tích vô hướng trong không gian các đa thức. Phương pháp tìm hàm xu thế trình bày ở đây tận dụng được ưu điểm là các thời điểm lấy số liệu thống kê cách đều nhau. Ngoài ra, phương pháp này có độ chính xác cao và đơn giản trong tính toán.

2. Tìm hàm xu thế nhờ tích vô hướng.

Gọi $x_0, x_1, \dots, x_n$ là các thời điểm lấy số liệu thống kê với các số liệu thống kê tương ứng là $y_0, y_1, \dots, y_n$. Bài toán đặt ra là tìm một đa thức $\hat{P}(x)$ bậc $k < n$ gần với các điểm $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ nhất theo nghĩa:

$$\begin{cases} \hat{P}(x_0) = y_0 + e_0 \\ \hat{P}(x_1) = y_1 + e_1 \\ \dots\dots\dots \\ \hat{P}(x_n) = y_n + e_n \end{cases} \quad (1)$$

và $\varepsilon = (e_0^2 + e_1^2 + \dots + e_n^2)^{1/2}$ đạt giá trị nhỏ nhất

Trong các giáo trình về kinh tế lượng có trình bày nhiều phương pháp giải quyết bài toán nêu trên (xem [1], [2]).

Ở đây, chúng tôi trình bày một cách giải dựa vào một tích vô hướng trong không gian các đa thức như sau:

Gọi $\mathcal{H}_n$ là không gian tất cả các đa thức bậc $\leq n$. Có định $x_0, x_1, \dots, x_n$. Với $p, q \in \mathcal{H}_n$, đặt:

$$\langle p, q \rangle = p(x_0) \cdot q(x_0) + p(x_1) \cdot q(x_1) + \dots + p(x_n) \cdot q(x_n) \quad (2)$$

Đây là một tích vô hướng trong không gian vectơ $\mathcal{H}_n$.

Tương ứng với tích vô hướng (2), ta có cách đo khoảng cách giữa hai đa thức $p, q \in \mathcal{H}_n$ xác định bởi:

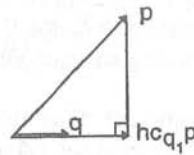
$$d(p, q) = \left\{ \sum_{i=0}^n [p(x_i) - q(x_i)]^2 \right\}^{1/2} \quad (3)$$

Với $p, q \in \mathcal{H}_n$, ta có hình chiếu của đa thức p xuống đa thức q xác định bởi (hình 1):

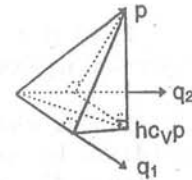
$$hc_{q,p} = \frac{\langle p, q \rangle}{\langle q, q \rangle} q = \frac{\sum_{i=0}^n p(x_i)q(x_i)}{\sum_{i=0}^n q^2(x_i)} q \quad (4)$$

Nếu q_1, q_2 là hai đa thức trực giao trong $\mathcal{H}_n$ theo nghĩa $\langle q_1, q_2 \rangle = 0$ thì hình chiếu của $p \in \mathcal{H}_n$ xuống không gian vectơ con V sinh bởi $\{q_1, q_2\}$ xác định như sau: (hình 2)

$$hc_{V,p} = hc_{q_1,p} + hc_{q_2,p} = \frac{\langle p, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 + \frac{\langle p, q_2 \rangle}{\langle q_2, q_2 \rangle} q_2 \quad (5)$$



Hình 1



Hình 2

Xét không gian vectơ con $\mathcal{H}_k$ của không gian $\mathcal{H}_n$ ($k < n$). Gọi $\{q_0, q_1, \dots, q_k\}$ là một cơ sở trực giao (gồm các đa thức đôi một trực giao) của $\mathcal{H}_k$. Ta có hình chiếu của đa thức $p \in \mathcal{H}_n$ xuống không gian con $\mathcal{H}_k$ xác định bởi:

$$hc_{\mathcal{H}_k,p} = \sum_{i=0}^k hc_{q_i,p} \quad (6)$$

Gọi $\{p_0, p_1, \dots, p_k\}$ là một cơ sở của không gian con $\mathcal{H}_k$. Từ cơ sở này, ta có thể xây dựng một cơ sở trực giao của $\mathcal{H}_k$ bằng các công thức sau: (xem [3])

$$\begin{cases} q_0 = p_0 \\ q_1 = p_1 - hc_{q_0,p_1} p_1 \\ \dots \\ q_k = p_k - \sum_{i=0}^{k-1} hc_{q_i,p_k} p_i \end{cases} \quad (7)$$

Bây giờ, ta trở lại bài toán (1). Ta có duy nhất một đa thức $\hat{p}$ bậc n thỏa:

$$\begin{cases} \hat{p}(x_0) = y_0 \\ \hat{p}(x_1) = y_1 \\ \dots \\ \hat{p}(x_n) = y_n \end{cases} \quad (8)$$

Từ (1) và (3), ta có: $\varepsilon = d(p, \hat{p})$

Như thế, cần tìm đa thức $\hat{p}$ bậc $k < n$ sao cho: $d(p, \hat{p})$ đạt giá trị nhỏ nhất.

Do tính chất của hình chiếu, ta có:

$$\hat{p} = hc_{\mathcal{H}_k,p} = \sum_{i=0}^k hc_{q_i,p} \quad (9)$$

Lưu ý rằng với tích vô hướng (2) việc tìm $hc_{q_i}p$ chỉ cần dùng các giá trị $p(x_0) = y_0, p(x_1) = y_1, \dots, p(x_n) = y_n$ mà không cần tìm p . Nói cách khác, với tích vô hướng (2), một đa thức $p \in P_n$ được đồng nhất với một vectơ $n+1$ chiều:

$$p(p(x_0), p(x_1), \dots, p(x_n)) \quad (10)$$

Để minh họa, chúng ta xét một ví dụ đơn giản như sau:

Ví dụ:

Tìm đa thức $\hat{p}$ bậc 2, gần nhất với các điểm dữ kiện: $(-2, 3), (-1, 5), (0, 5), (1, 4), (2, 3)$

Ta xét không gian vectơ $\mathcal{H}_4$ (do có 5 điểm dữ kiện), với tích vô hướng:

$$\langle p, q \rangle = p(-2) \cdot q(-2) + p(-1) \cdot q(-1) + p(0) \cdot q(0) + p(1) \cdot q(1) + p(2) \cdot q(2)$$

Ta có: $p_0 = 1, p_1 = x, p_2 = x^2$ là các đa thức độc lập tuyến tính; lập thành một cơ sở của không gian con $\mathcal{H}_2$ của $\mathcal{H}_4$.

Điều này chứng minh dễ dàng bằng cách đồng nhất một đa thức bậc ≤ 4 với một vectơ 5 chiều gồm các giá trị của đa thức đó tại $x_0 = -2, x_1 = -1, x_2 = 0, x_3 = 1, x_4 = 2$. Ta có:

$$\begin{aligned} p &= (3, 5, 5, 4, 3) \\ p_0 &= (1, 1, 1, 1, 1) \\ p_1 &= (-2, -1, 0, 1, 2) \\ p_2 &= (4, 1, 0, 1, 4) \end{aligned}$$

Bây giờ, ta xây dựng một cơ sở trực giao $\{q_0, q_1, q_2\}$ cho $\mathcal{H}_2$ từ $\{p_0, p_1, p_2\}$. Do công thức (7), ta có:

$$q_0 = p_0 = 1$$

$$q_1 = p_1 - hc_{q_0}p_1 = x - \frac{\langle p_1, q_0 \rangle}{\langle q_0, q_0 \rangle} q_0 = x - \frac{0}{5} \cdot 1 = x$$

$$\begin{aligned} q_2 &= p_2 - hc_{q_0}p_2 - hc_{q_1}p_2 = x^2 - \frac{\langle p_2, q_0 \rangle}{\langle q_0, q_0 \rangle} q_0 - \frac{\langle p_2, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 \\ &= x^2 - \frac{10}{5} \cdot 1 - \frac{0}{10} x = x^2 - 2 \end{aligned}$$

$$\text{Suy ra: } \hat{p} = hc_{q_0}p + hc_{q_1}p + hc_{q_2}p$$

$$= \frac{\langle p, q_0 \rangle}{\langle q_0, q_0 \rangle} q_0 + \frac{\langle p, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 + \frac{\langle p, q_2 \rangle}{\langle q_2, q_2 \rangle} q_2$$

$$\begin{aligned} \text{với } q_0 &= 1 \equiv (1, 1, 1, 1, 1) \\ q_1 &= x \equiv (-2, -1, 0, 1, 2) \\ q_2 &= x^2 - 2 \equiv (2, -1, -2, -1, 2) \\ \langle p, q_0 \rangle &= 20, \langle q_0, q_0 \rangle = 5, \\ \langle p, q_1 \rangle &= -1, \langle q_1, q_1 \rangle = 10 \\ \langle p, q_2 \rangle &= -7, \langle q_2, q_2 \rangle = 14 \\ \text{Vậy: } \hat{p} &= 4 - \frac{1}{10}x - \frac{1}{2}(x^2 - 2) \\ &= 5 - \frac{1}{10}x - \frac{1}{2}x^2 \end{aligned}$$

là hàm xu thế bậc 2 xấp xỉ tốt nhất 5 điểm dữ kiện trên.

Ta có thể tính sai số ϵ như sau:

$$\hat{p} \equiv \left(\frac{16}{5}, \frac{23}{5}, \frac{22}{5}, \frac{14}{5}, \frac{14}{5} \right)$$

$$p = (3, 5, 5, 4, 3)$$

Ta có:

$$\begin{aligned} \epsilon = d(p, \hat{p}) &= \left[\left(\frac{16}{5} - 3 \right)^2 + \left(\frac{23}{5} - 5 \right)^2 + \left(\frac{22}{5} - 5 \right)^2 + \left(\frac{14}{5} - 4 \right)^2 + \left(\frac{14}{5} - 3 \right)^2 \right]^{1/2} \\ &= \frac{\sqrt{10}}{5} \end{aligned}$$

Bây giờ, thử tìm hàm xu thế bậc 2 phản ánh gần đúng nhất các số liệu tiêu dùng cá nhân ở Mỹ vào các năm: 1970 là 672,1 USD, 1971 là 696,8 USD, 1972 là 737,1 USD, 1973 là 767,9 USD, 1974 là 762,8 USD (xem [4], trang 143). Đặt: 1970 = -2, 1971 = -1, 1972 = 0, 1973 = 1, 1974 = 2, ta có:

$$p \equiv (672,1; 696,8; 731,1; 767,9; 762,8)$$

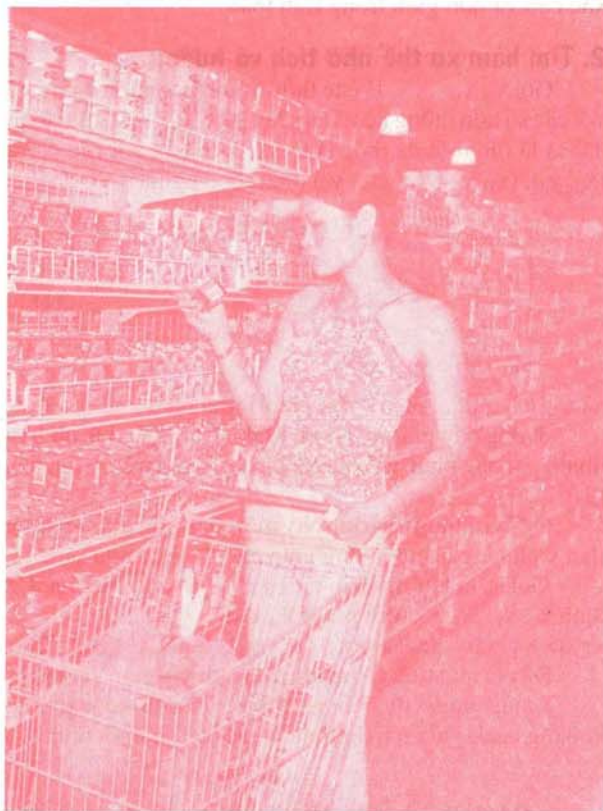
$$\text{Và: } \langle p, q_0 \rangle = 3.636,7; \langle p, q_1 \rangle = 252,5; \langle p, q_2 \rangle = -69,1;$$

$$\begin{aligned} \text{Suy ra: } \hat{p} &= \frac{3.636,7}{5} + \frac{252,5}{10}x - \frac{69,1}{10}(x^2 - 2) \\ &= 737,21143 + 25,25x - 4,9357x^2 \end{aligned}$$

Chú thích: Nếu các điểm dữ kiện là số lẻ, bằng một phép đổi biến, ta có thể giả sử $x_0, x_1, \dots, x_n$ cách đều hai bên số 0 như trong ví dụ trên. Khi đó, việc tính các tích vô hướng rất đơn giản. Ngoài ra, nếu dãy số liệu càng dài thì hàm xu thế càng chính xác hơn ■

TÀI LIỆU THAM KHẢO

- [1] Ameriya, T., Advanced Econometrics, Cambridge, Mass.: Harvard University Press, 1985.
- [2] Schmidt, P., Econometrics, Newyork: Marcel Dekker, 1976
- [3] Rudin, W., Real and Complex Analysis, MacGraw-Hill, 1974
- [4] Green, W.H, Econometrics Analysis, Mac Millan 1991



Ảnh: Hòa Tấn