

GIỚI THIỆU PHƯƠNG PHÁP ENTRÔPI CỰC ĐẠI (MEM)

TRONG TÍNH ƯỚC LƯỢNG PHỔ

Nguyễn Thuyết

Phòng Khoa học kỹ thuật

Trong thực tế, chúng ta tính ước lượng phổ bằng những chuỗi số liệu có độ dài hữu hạn của các quá trình tự nhiên. Điều đó gây hiệu ứng thẩm lậu năng lượng, làm sai lệch ước lượng phổ. Và việc tính còn đòi hỏi phải có chuỗi số liệu dài ngay mà nhiều khi kho số liệu có sẵn không đáp ứng được. Vì thế, về nguyên tắc, càng kéo dài được chuỗi số liệu càng có được những ước lượng phổ tốt hơn. Việc tìm kiếm cách có được những ước lượng phổ tốt hơn trong điều kiện bị hạn chế bởi chuỗi số liệu ngắn ngay đã dẫn đến việc ứng dụng phương pháp MEM để tính ước lượng phổ.

Về nguyên lý, việc tính ước lượng phổ bằng MEM là ngoại suy trường các giá trị của hàm tự hiệp biến (hoặc hệ số tự tương quan) đến bước trễ vô hạn bằng cách sao cho entropi (độ đo thông tin) được cực đại hóa.

Trong thực tế, việc tính ước lượng phổ bằng MEM lại được thực hiện bằng cách tiên đoán chuỗi số liệu đến thời gian vô hạn bằng cách làm thích hợp chuỗi tự hồi qui AR, là thể hiện của chuỗi số liệu theo thời gian ấy.

Và như vậy, theo nguyên tắc thì ước lượng phổ tính được bằng MEM gần với phổ thật của quá trình hơn là ước lượng phổ tính được bằng phương pháp thường tính. Ưu điểm nổi bật của MEM là cho một ước lượng phổ mịn hơn so với các phương pháp khác, dù rằng chỉ sử dụng chuỗi số liệu có độ dài ngắn hơn.

Phương pháp MEM đầu tiên do Burg đề xuất năm 1967, sau đó được các tác giả Ulrych, Smylie, Linton, Barnand, Edward, Jenkins và Watts, Van den Box nghiên cứu phát triển thêm về mặt lý thuyết; Andersen đã đưa ra một sơ đồ đệ qui tổng quát để tính các sai số tiên đoán (các hệ số AR), và như là Akaike đã đề xuất một tiêu chuẩn để xác định độ dài thích hợp của chuỗi các hệ số AR. Nhiều tác giả đã nghiên cứu ứng dụng MEM vào việc nghiên cứu các hiện tượng địa vật lý và cả trong lãnh vực thiên văn như nghiên cứu chuyển động Chandler. Trong đó phải kể đến các tác giả Wells, Chinnery, Curie và gần đây là Hayashi đã ứng dụng MEM để tính phổ hai thứ nguyên (không và thời gian) và phổ giao hai thứ nguyên (không - thời gian) của trường các yếu tố tự nhiên.

Trước hết, chúng ta làm quen với một số khái niệm toán học ít gặp trong chuyên môn của chúng ta.

Entropi.

Xét một hệ thống trong đó có thể xảy ra m_i kiểu khác nhau của một hiện tượng i , với xác suất xuất hiện là p_i . Nếu tất cả các p_i đều bằng nhau, ta không

có thông tin đặc biệt gì về hệ thống. Nhưng nếu ta có thể xác định một p_1 đặc biệt thì ta bảo có được một lượng tin nào đó về hệ thống. Như vậy xác suất xuất hiện p_1 của một hiện tượng có liên quan với thông tin. Ta biểu diễn mối quan hệ này bằng biểu thức :

$$I = k \ln \frac{1}{p_1} \quad (1)$$

Nếu dùng lôga cơ số 2 ? thì hằng số k ? bằng 1 và biểu thức thành :

$$I = \log_2 \frac{1}{p_1} \quad (2)$$

Giả sử ta quan sát hệ thống trên trong một thời gian dài T , và thấy xảy ra $p_1 T$ các kiểu m_1 , $p_2 T$ các kiểu m_2 , ..., $p_n T$ các kiểu m_n . Shannon (1948) đưa ra khái niệm thông tin trung bình trong một khoảng thời gian và gọi bằng thuật ngữ entrôpi là :

$$H = -k \sum_{i=1}^N p_i \ln p_i = - \sum_{i=1}^N p_i \log_2 p_i \quad (3)$$

Vì vậy một cách thô thiển, entrôpi H được coi là độ đo thông tin. Jaynes đã nêu ra nguyên tắc :

"Sự phân bố xác suất tiên nghiệm mô tả thông tin đã biết nhưng không bị ràng buộc một cách tối đa với thông tin chưa biết là sự phân bố với entrôpi cực đại".

Ma trận tương quan.

Trong phân tích phổ để xử lý các chuỗi rời rạc theo thời gian, các phép tính thường được biểu thị một cách chặt chẽ dưới dạng ma trận. Vì thế các giá trị của hàm tương quan cũng được biểu diễn dưới dạng ma trận.

Trong một chuỗi thời gian, không phải thời gian được dùng làm đối số mà lại là số thứ tự i của các giá trị của chuỗi. Từ đó các giá trị của hàm tương quan (ở đây là hàm tự hiệp biến) có thể được đánh dấu bằng 2 chỉ số cấu thành từ số thứ tự của hai giá trị của chuỗi thời gian :

$$R_{ij} = E \{ x_i x_j \} \quad (4)$$

$E \{ \}$ là ký hiệu của kỳ vọng toán.

Đối với một chuỗi dừng, giá trị của R_{ij} chỉ phụ thuộc vào môđun của hiệu các số thứ tự $|j - i|$, vì thế

$$R_{ij} = R_{ji} \text{ và } R_{ij} = R_{i+n, j+n} \quad (5)$$

Các giá trị của R_{ij} được sắp xếp thành ma trận như sau :

$$\begin{bmatrix} R_{11} & R_{12} & \dots & R_{1p} \\ R_{21} & R_{22} & \dots & R_{2p} \\ \dots & \dots & \dots & \dots \\ R_{p1} & R_{p2} & \dots & R_{pp} \end{bmatrix}$$

trong đó : $R_{11} = R_{22} = \dots = R_{pp}$ (6)

và $R_{21} = R_{12} = R_{32} = R_{31} = \dots = R_{p-1 p} = R_{p p-1}$

Cụ thể có một chuỗi giá trị hàm tương quan $R_0, R_1, \dots, R_p$ (gồm có $p+1$ giá trị), thì có thể biểu diễn thành ma trận tương quan như sau :

$$\begin{bmatrix} R_0 & R_1 & R_2 & \dots & R_p \\ R_1 & R_0 & R_1 & \dots & R_{p-1} \\ R_2 & R_1 & R_0 & \dots & R_{p-2} \\ \dots & \dots & \dots & \dots & \dots \\ R_p & R_{p-1} & \dots & \dots & R_0 \end{bmatrix} \quad (6')$$

Mô hình tự hồi quy.

Từ năm 1927, Yule đã chỉ ra rằng một quá trình có thể được mô tả bằng một mô hình lọc tuyến tính có dạng :

$$x_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \quad (7)$$

Từ tư tưởng trên, thông qua một số biến đổi, người ta có thể biểu diễn một giá trị hiện tại của một quá trình thông qua một số hữu hạn các giá trị trước nó:

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_p x_{t-p} + \xi(t) \quad (8)$$

trong đó $\{a_j\}$ ($j = 1, 2, \dots, p$) là tập các hệ số trọng lượng; p là bậc của mô hình; $\xi(t)$ là một giá trị của một chuỗi không tương quan gọi là nhiễu trắng. Cách thể hiện như vậy gọi là một mô hình tự hồi quy (AR).

Một chuỗi mà trong đó một giá trị hiện tại được thể hiện bằng một mô hình AR gọi là một chuỗi tự hồi quy.

Sự tương ứng giữa quá trình AR với MEM.

1. Ứng dụng khái niệm entropy vào phân tích phổ, Smylie đã xây dựng biểu thức về mối quan hệ giữa entropy với một độ phổ $E(f)$ (phổ năng lượng) của một quá trình Gauss dừng :

$$H = \frac{1}{4f_N} \int_{-f_N}^{f_N} \log E(f) df \quad (9)$$

trong đó f_N là tần số Nyquist (tần số cao nhất ứng với bước thời gian Δt của quá

trình, $f_N = \frac{1}{2\Delta t}$).

Ông cũng đã nêu ra biểu thức:

$$H = -\frac{1}{2} \log (\det [C_p]) \quad (10)$$

trong đó $[C_p]$ là ma trận tương quan được cho bởi (6') $\det []$ là ký hiệu của định thức của ma trận. Từ chuỗi số liệu ta tính được $p+1$ giá trị của $[C_p]$ từ R_0 đến R_p , và tư tưởng của MEM là ngoại suy các giá trị $R_{p+1}, R_{p+2}, \dots$ v.v. với điều kiện cực đại hóa entropi từng bước.

Điều kiện để cho entropi cực đại khi ngoại suy R_{p+1} từ $[C_p]$ là:

$$\det \begin{bmatrix} R_1 & R_0 & R_1 & \dots & R_{p-1} \\ R_2 & R_1 & R_0 & \dots & R_{p-2} \\ \dots & \dots & \dots & \dots & \dots \\ R_{p+1} & R_p & \dots & \dots & R_1 \end{bmatrix} = 0 \quad (11)$$

2. Bây giờ trở lại xét một quá trình AR được cho bởi (8), viết lại như sau:

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_p x_{t-p} + \xi(t) \quad (8)$$

Nhân 2 vế của (8) với x_{t-k} , rồi lấy giá trị kỳ vọng của chúng, và vì $E(x_{t-k} \xi(t)) = 0$ với $k > 0$, ta có:

$$R_k = a_1 R_{k-1} + a_2 R_{k-2} + \dots + a_p R_{k-p} \quad (12)$$

Trong (12) thay k lần lượt bằng 1, 2, $\dots$ $p+1$, ta sẽ được hệ phương trình Yule - Walker:

$$\left. \begin{aligned} R_1 - a_1 R_0 - a_2 R_1 - \dots - a_p R_{p-1} &= 0 \\ R_2 - a_1 R_1 - a_2 R_0 - \dots - a_p R_{p-2} &= 0 \\ \dots & \\ R_{p+1} - a_1 R_p - \dots - a_p R_1 &= 0 \end{aligned} \right\} \quad (13)$$

Trong hệ phương trình (13), $p+1$ giá trị đầu $R_0, R_1, \dots, R_p$ của hàm tương quan đã biết, còn giá trị R_{p+1} có thể được xác định bằng cách giải:

$$\det \begin{bmatrix} R_1 & R_0 & R_1 & \dots & R_{p-1} \\ R_2 & R_1 & R_0 & \dots & R_{p-2} \\ \dots & \dots & \dots & \dots & \dots \\ R_{p+1} & R_p & \dots & \dots & R_1 \end{bmatrix} = 0 \quad (14)$$

So sánh (14) với (11) cho thấy sẽ được cùng một giá trị của R_{p+1} trong 2 cách tính. Sự phù hợp này cũng đúng với mọi giá trị ngoại suy sau R_{p+1} của hàm tương quan. Vì thế có thể kết luận rằng sự phân tích phổ MEM tương đương với sự làm thích hợp một mô hình AR để ngoại suy cho một quá trình ngẫu nhiên.

Tiêu chuẩn FPE của Akaike.

Từ thực nghiệm người ta thấy rằng độ dài của bộ lọc (một cách tương đương là bậc của mô hình AR) giữ một vai trò trọng yếu trong phép phân tích phổ MEM. Độ dài ngắn quá sẽ dẫn đến việc làm trơn ước lượng phổ quá nhiều, xóa bỏ lợi thế về khả năng phân giải của MEM; còn độ dài lớn quá sẽ đưa vào phổ những chi tiết giả.

Nghiên cứu vấn đề này, Akaike đã nhận thấy phương sai σ^2 của nhiễu trắng (chuỗi $\xi(t)$) trong mô hình AR rất nhạy cảm với độ dài của bộ lọc. Khi độ dài n của y tăng thì phương sai σ^2 giảm rất nhanh, sau đó chậm dần đến một cực tiểu, rồi lại tăng lên. Giá trị p của bậc của mô hình AR làm cho σ^2 cực tiểu được chọn làm độ dài của bộ lọc. Akaike đưa khái niệm sai số tiên đoán quân phương FPE, tính được bằng biểu thức:

$$FPE = E \left[(x_t - \hat{x}_t)^2 \right] \quad (15)$$

trong đó x_t là ước lượng của giá trị x_t của chuỗi số liệu theo mô hình AR. Độ dài lọc tốt nhất là làm cho FPE cực tiểu. Akaike đã xây dựng biểu thức quan hệ giữa FPE với độ dài N của chuỗi số liệu, độ dài tối ưu P của bộ lọc và phương sai σ^2 .

$$FPE(p) = \left(1 + \frac{p+1}{N} \right) \sigma_{(p)}^2 \quad (16)$$

Gần đây, Akaike đưa ra hướng mới là phải cực tiểu hóa một tổng, gọi là tiêu chuẩn thông tin:

$$Y = \ln \sigma_{(p)}^2 + \frac{2P}{T} \quad (17)$$

...

Như vậy, thực chất của vấn đề có thể hiểu là tính ước lượng phổ MEM là tính ước lượng phổ theo tập hệ số $\{a_p\}$ của mô hình AR được dùng để tiên đoán giá trị x_p (của chuỗi số liệu) theo $p-1$ giá trị trước của chuỗi. Công thức để tính ước lượng phổ MEM như sau:

$$E(f) = \frac{\sigma_{(p)}^2}{\left| 1 - \sum_{j=a}^p a_j \exp(-i2\pi f j) \right|^2} \quad (18)$$

Tập hệ số a_p tính được từ hệ phương trình Yule-Walker, hoặc tính theo sơ đồ Burg, bằng công thức đệ quy của Andersen:

$$a_m^{(p)} = a_m^{(p-1)} + a_p^{(p)} a_{p-m}^{(p-1)}, m = 1, \dots, p. \quad (19)$$

Chỉ số ở trên (p) là chỉ bậc của mô hình AR. Hệ số $a_p^{(p)}$ tính được từ công thức :

$$\frac{\partial S(a_p^{(p)})}{\partial a_p^{(p)}} = 0 \quad (20)$$

Đây là điều kiện để cực tiểu hóa hàm tổng dư bình phương $S(a_p^{(p)})$, mà :

$$S(a_p^{(p)}) = \frac{1}{2(N-m)} \sum_{j=m+1}^N \left\{ \left(x_j - \sum_{k=1}^p a_k^{(p)} x_{j-k} \right)^2 + \left(x_{j-p} - \sum_{k=1}^p a_k^{(p)} x_{j-m+k} \right)^2 \right\} \quad (21)$$

Phương sai $\sigma_{(p)}^2$ cũng được tính theo thuật toán đệ qui :

$$\sigma_{(p)}^2 = \sigma_{(p-1)}^2 (1 - [a_p^{(p)}]^2) \quad (22)$$

Ước lượng phổ MEM, cũng có thể tính được theo công thức :

$$E(f) = \frac{2 \sigma_{(p)}^2}{1 - \sum_{j=0}^p q_j \exp(-i2\pi f j)} \quad (23)$$

Trong đó tập $\{q_j\}$ là các hệ số tương quan của bộ lọc trung bình. Ma trận của hệ số tương quan q_j này là nghịch đảo của ma trận tương quan của chuỗi số liệu :

$$[q_p] = [c_p]^{-1} \quad (24)$$

Trong bộ chương trình tính phổ (bằng ngôn ngữ Fortran), đang được xây dựng, chúng tôi bổ sung thêm một chương trình con tính phổ MEM./.

Tài liệu tham khảo

- Tad J. Ulrych và Thomas Bishop. - Maximum entropy spectral analysis and Autoregressive decomposition. Tạp chí Geophysics and Space Physics. Vol 13. N°1 - 1975.
- K.V. Coniev. - Phân tích phổ trường hải văn ngẫu nhiên (tiếng Nga). NXB KTTV Leningrát - 1981.
- Trương Quang Hào. - Ứng dụng MEM để phân tích trường địa từ trên prôfom dài (tiếng Nga). Acta geophysica Polonica. Vol. XXIX - N°2.
- Nguyễn Thuyết. - Phân tích phổ. một công cụ nghiên cứu tính tế trong điều tra cơ bản. Nội san KTTV số 2-1983.