MÔ HÌNH TOÁN HỌC CỦA TỔ HỢP NGUYÊN TỐ CHỈ THỊ QUẶNG TRONG TRƯỜNG ĐỊA HOÁ THỨ SINH VÙNG TÂY BẮC VIỆT NAM

Đặng Mai

Khoa Địa Chất, Trường Đại học Khoa học Tự nhiên, ĐHQG Hà Nội

1. Mở đầu

Các tổ hợp nguyên tố chỉ thị (THNTCT) quặng trong trường địa hoá thứ sinh có ý nghĩa quan trọng trong việc đánh giá kiểu quặng hoá và tính chất sinh khoáng của khu vực. Mặt khác, đối với công tác tìm kiếm địa hoá, có thể dựa vào các tổ hợp đó để lựa chọn phương pháp và chỉ tiêu phân tích hợp lý. Tổ hợp nguyên tố chỉ thị thường được xác định bằng cách chồng ghép các bản đồ dị thường địa hoá hoặc phân tích tương quan. Mỗi phương pháp đều có những ưu thế và hạn chế nhất định. Trong bài báo này, lần đầu tiên ở Việt Nam, trong lĩnh vực này sẽ ứng dụng thuật toán phân tích chùm (cluster analyse) và phần mềm SPSS FOR WINDOWS 10.0 để xác định THNTCT.

Cơ sở số liệu gồm 6000 mẫu kim lượng vùng Tây Bắc Việt Nam do các đoàn địa chất thu thập. Các mẫu được phân tích bằng phương pháp quang phổ bán định lượng toàn phần bao gồm các nguyên tố Ba, Ti, Cr, Co, Ni, Mo, W, Sn, Bi, Cu, Pb, Zn, Nb, Li, Ce, La, Zr.

2. Cơ sở lý thuyết của phương pháp

Phương pháp phân tích chùm là phương pháp ghép nhóm dựa vào độ đo (measure) khoảng cách của các đối tượng, mà trong trường hợp này là các nguyên tố trong trường địa hoá thứ sinh. Kết quả của việc ghép nhóm là đưa ra biểu đồ phân loại dạng cành cây (dendrogram). Dưới đây, sẽ nêu khái quát thuật toán phân loại và các độ đo đã được sử dụng trong bài báo này.

Thuật toán phân loại

Để thiết lập các mô hình THNTCT, áp dụng ba thuật toán phân loại sau:

- Lân cận gần nhất (nearest neighbor).
- Trung vị (median clustring).
- Trọng tâm (centroid clustering).

Các thuật toán này đều dựa trên ma trận khoảng cách của các đối tượng và sử dụng thủ tục hợp nhất liên tiếp để ghép nhóm. Đầu tiên, coi n đối tượng cần phân loại là n nhóm, sau đó ghép 2 đối tượng có khoảng cách bé nhất vào thành một tập. Sau bước này, tập hợp ban đầu được phân thành n - 1 tập con, trong đó có một tập

con gồm hai đối tượng, còn các tập khác gồm 1 đối tượng. Tiếp theo, lại so sánh từng cặp tập con khác nhau và gộp hai tập con nào có khoảng cách bé nhất lại với nhau thành một tập con mới. Thuật toán tiếp tục như vậy cho đến khi tất cả đối tượng ghép thành một nhóm. Số tập con ở bước trước đó là số nhóm cần phân chia. Trong thủ tục này, từ bước thứ hai trở đi việc so sánh khoảng cách giữa hai nhóm gồm ba trường hợp.

- Khoảng cách giữa đối tượng với đối tượng.
- Khoảng cách giữa đối tượng với một nhóm đối tượng.
- Khoảng cách giữa các nhóm đối tượng với nhau.

Ba thuật toán nêu trên khác nhau ở phương pháp xác định khoảng cách của các nhóm đối tượng sau mỗi bước ghép nhóm.

Trong thuật toán lân cận gần nhất, khoảng cách giữa hai nhóm được lấy bằng khoảng cách giữa hai đối tượng gần nhất trong hai nhóm đó.

Trong thuật toán trọng tâm, khoảng cách giữa hai nhóm là khoảng cách giữa hai điểm trung bình của tất cả các đối tượng trong nhóm.

Trong thuật toán trung vị, hai nhóm có trung bình trọng bằng nhau sẽ được ghép thành nhóm mới.

Các độ đo

Năm độ đo sau đây đã được áp dụng:

1. Khoảng cách Ocolit:

$$D(x, y) = \sqrt{\sum_i (x_i - y_i)^2}. \quad (1)$$

2. Khoảng cách khi bình phương:

$$D(x, y) = \sqrt{\sum_i \frac{(x_i - E(x_i))^2}{E(x_i)} + \sum_i \frac{(y_i - E(y_i))^2}{E(y_i)}}. \quad (2)$$

3. Khoảng cách Trêbusev:

$$D(x,y) = \mathrm{Max}_i \mid x_i - y_i \mid. \quad (3)$$

4. Khoảng cách Cosine:

$$D(x, y) = \frac{\left(\sum x_i y_i\right)}{\sqrt{\left(\sum_i x_i^2\right)\left(\sum_i y_i^2\right)}}. \quad (4)$$

5. Độ đo Pearson:

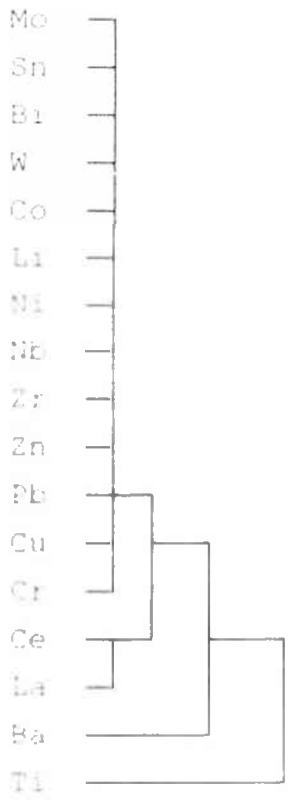
$$D(x, y) = \frac{\sum_{i} Z_{xi} Z_{yi}}{N - 1}, \quad (5)$$

Trong các biểu thức trên:

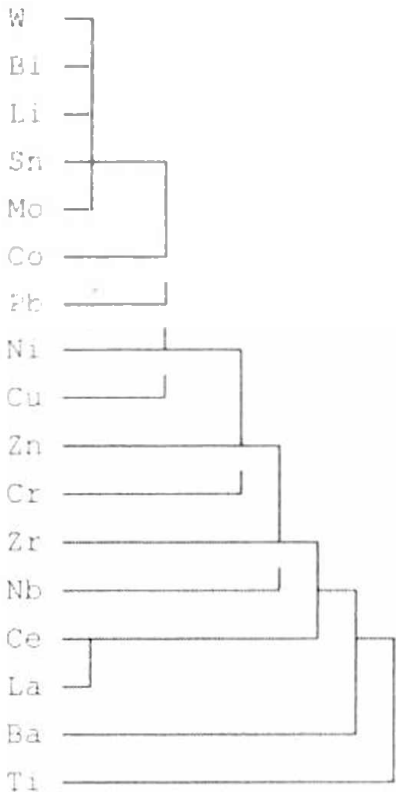
- x_i và y_i là giá trị của thuộc tính i trên đối tượng x và y tương ứng.
- Z_{xi} - giá trị chuẩn hoá của x_i, Z_{yi} - của y_i
- $E(x_i)$, $E(y_i)$ - kỳ vọng của x_i và y_i tương ứng
- N là tổng số thuộc tính.

3. Mô hình tổ hợp nguyên tố chỉ thị

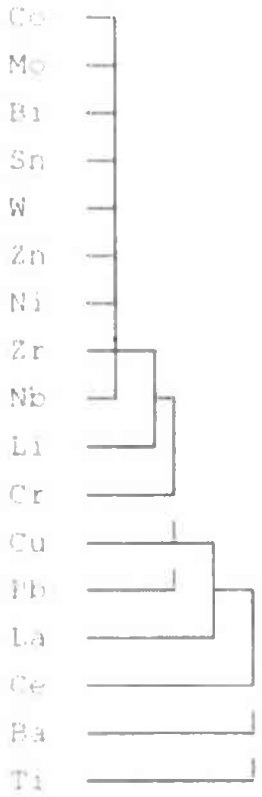
Mô hình THNTCT được đưa ra trên các hình từ 1 đến 8. Đó là các dendrogram được xử lý và thiết lập bằng chương trình SPSS 10.0 theo các thuật toán nêu trên.



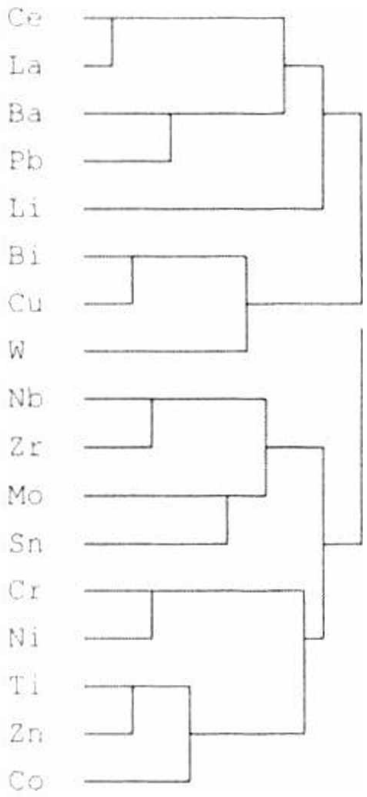
Hình 1. Mô hình Ocolit - thuật toán trọng tâm



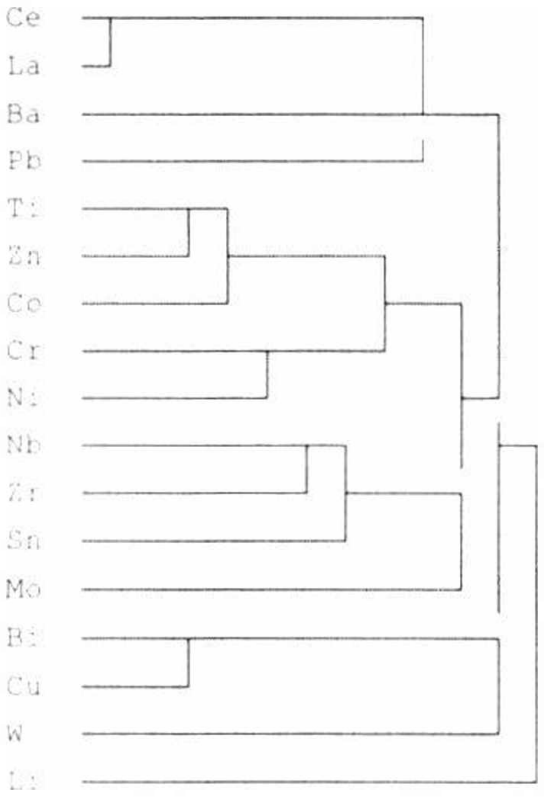
Hình 2. Mô hình Khi bình phương - thuật toán trọng tâm



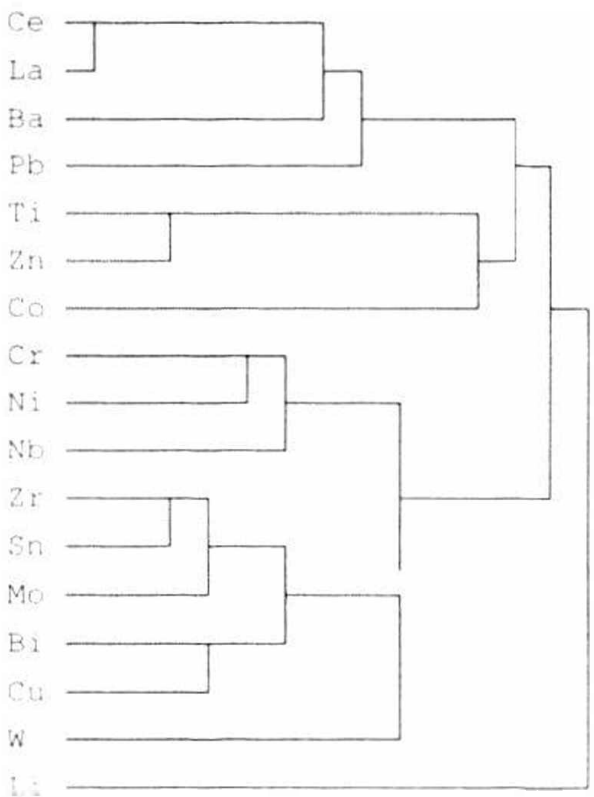
Hình 3. Mô hình Trebưsev -thuật toán trọng tâm



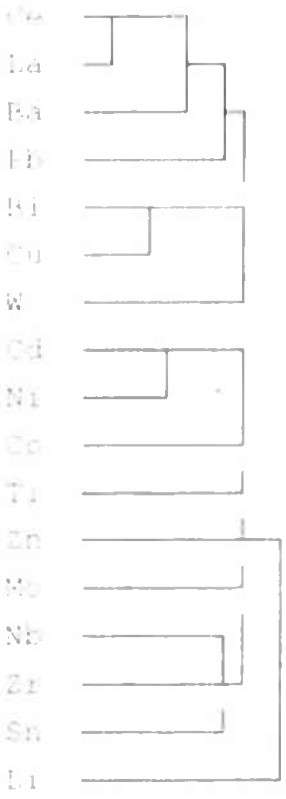
Hình 4. Mô hình Cosin- thuật toán trọng tâm



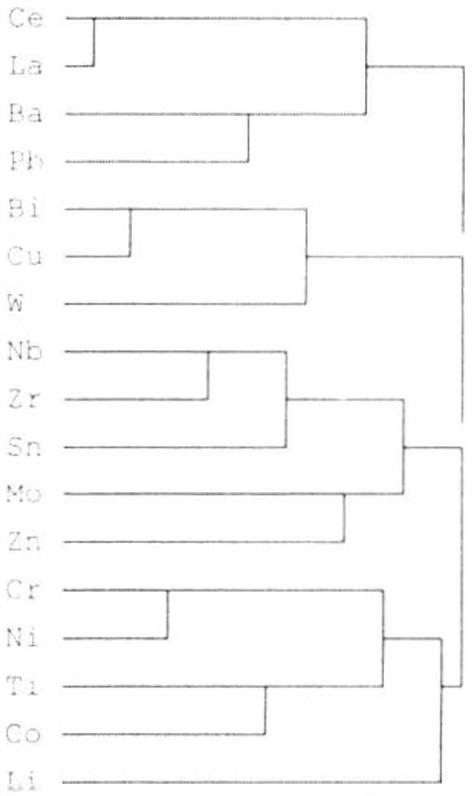
Hình 5. Mô hình Cosine - thuật toán trung vị



Hình 6. Mô hình Cosine - thuật toán lân cận gần nhất



Hình 7. Mô hình Pearson - thuật toán trung vị



Hình 8. Mô hình Pearson - thuật toán trọng tâm

4. Thảo luận kết quả

Đendrogram lập theo khoảng cách Ơcơlit gồm hai nhánh (hình1), trong đó nhánh thứ nhất chỉ có titan, nhánh thứ hai chứa tất cả các nguyên tố còn lại. Đến lượt mình, nhánh thứ hai lại được tách thành hai nhánh con: bari và các nguyên tố khác v.v... Rõ ràng rằng mô hình này không phản ánh đúng đặc điểm địa hoá của khu vực nghiên cứu.

Mô hình lập theo khoảng cách khi bình phương cũng có những nét tương tự như mô hình trên hình 2. Ở đây, xuất hiện tổ hợp W - Sn - Mo - Li - Bi tương đối rõ rệt, còn các nguyên tố khác như Pb, Zn, Cu... không nằm trong một tổ hợp nào. Mô hình này cũng không hợp lý.

Trên hình 3 ta thấy có ba nhánh độc lập, đó là Ti, Ba, Ce và các nguyên tố còn lại. Nhánh thứ tư này lại phân ra các nhánh phụ bao gồm Co-Mo-Bi-Sn-W-Zn-Nb, Cr-Cu-Pb và La. Các tổ hợp này không phù hợp với tính chất địa hoá của các nguyên tố. Như vậy, độ đo Trêbưsev không thích hợp cho việc phân nhóm các tổ hợp nguyên tố chỉ thị trong trường địa hoá thứ sinh.

nhau: trọi Khoảng cách Cosine được thử nghiệm theo ba thuật toán khác nhau: trọng tâm của tâm, trung vị và lân cận gần nhất. Trên mô hình theo thuật toán trọng tâm của độ Ce, La, Bio này xuất hiện ba nhánh lớn. Nhánh thứ nhất bao gồm năm nguyên tố Ce, La, Ba, còn lại (N'b, Li. Nhánh thứ hai bao gồm ba nguyên tố Bi, Cu, W. Các nguyên tố còn lại (Nb, ong khi đr, Mo, Sn, Cr, Ni, Ti, Zn, Co) thuộc vào nhánh thứ ba (hình 4). Trong khi đó, h thứ nhẫendrogram theo phương pháp trung vị chỉ gồm hai nhánh: Li là nhánh thứ nhất, a thành lớn nhánh kia gồm 16 nguyên tố (hình 5). Nhánh thứ hai này lại chia thành ba gồm Ti, Zhánh con. Nhánh con thứ nhất gồm Ce, La, Ba, Pb. Nhánh con thứ hai gồm Ti, Zn, nh thứ bo, Cr, Ni, Nb, Zr, Sn, Mo. Các nguyên tố Bi, Cu, W thuộc nhánh thứ ba. ơng tự. Lidendrogram lập theo thuật toán lân cận gần nhất cũng có những nét tương tự. Liti nhánh phụ ẫn tách thành một nhánh độc lập, nhưng tổ hợp nguyên tố trong các nhánh phụ có tổ hợp thhiều thay đổi. Tổ hợp thứ nhất bao gồm Ce, La, Ba, Pb, Ti, Zn, Co và tổ hợp thứ hình Cosai gồm Cr, Ni, Nb, Zr, Sn, Mo, Bi, Cu, W. Như vậy, các tổ hợp theo mô hình Cosin hông thật rõ ràng và chưa phản ánh đúng thực tế.

như các n Chuyển sang các dendrogram lập theo độ đo Pearson. Tương tự như các mô ông cân đoinh Cosine, trong thuật toán trung vị, cây phân loại gồm hai nhánh không cân đối. Pb, Bi, CMột nhánh là Li, còn nhánh kia hợp thành bởi các nguyên tố Ce, La, Ba, Pb, Bi, Cu, ai này đưV, Cd, Ni, Co, Ti, Zn, Mo, Nb, Zr, Sn. Các nguyên tố trong nhóm thứ hai này được o và Nb-Zếp vào bốn phụ nhóm. Đó là Ce-La-Ba-Pb, Bi-Cu-W, Cd-Ni-Co-Ti-Zn-Mo và Nb-Zr- oại này kiên (hình 7). Nếu không xét trường hợp đứng riêng của Li, cây phân loại này khá hù hợp với thực tế.

tách thàn Đendrogram theo thuật toán trọng tâm của độ đo Pearson được tách thành La, Ba, Pốn nhánh rõ rệt (hình 8). Các thành viên trên nhánh thứ nhất là Ce, La, Ba, Pb. hứ hai gồ Đó là một tổ hợp đặc trưng cho quặng nhiệt dịch chứa đất hiếm. Nhánh thứ hai gồm c trưng cBi, Cu và W, tương ứng tổ hợp nhiệt dịch nhiệt độ cao. Nhánh thứ ba đặc trưng cho uối cùng, ác nguyên tố liên quan đến magma axit như Nb, Zr, Sn, Mo, Zn. Và cuối cùng, tổ hợp chỉ thợp nguyên tố liên quan đến magma bazơ tạo nên nhánh thứ tư. Các tổ hợp chỉ thị guyên tố thể hiện trên dendrogram này phản ánh đúng đặc điểm địa hoá các nguyên tố và ặc điểm kim sinh của khu vực nghiên cứu.

. Kết luận

iu: Từ những điều trình bày trên đây, có thể rút ra hai kết luận như sau:

Pb -La- Ce 1. Trong vùng nghiên cứu tồn tại bốn tổ hợp nguyên tố chỉ thị là Pb -La- Ce - a, Bi - Cu - W, Nb - Zr - Sn - Mo và, Cr - Ni - Ti - Co - Li.

chia các t 2. Độ đo Pearson với thuật toán trọng tâm thích hợp cho việc phân chia các tổ ợp nguyên tố chỉ thị trong trường địa hoá thứ sinh.

ài liệu tham khảo

Universit Albarède F., *Introduction to Geochemical Modeling*, Cambridge University Press 1995, 543 pp

2. Pan G. & D. P. Harris, Quantitative analysis of anomalous sources an geochemical signatures in the Walker Lake quadrangle of Nevada an California, *Geo.Exp.*, Elsevier, Amsterdam, V.38, N° 3 (1990), pp 299 -321.

3. Nguyễn Văn Liêu, Nguyễn Đình Cử, Nguyễn Quốc Anh, *SPSS - Ứng dụng phá tích dữ liệu trong quản trị kinh doanh và khoa học tự nhiên - xã hội*, NXB Gia thông Vận tải, Hà Nội, 2000, 296 trang.

4. Đặng Mai, Hoàng Minh, Đỗ Văn Phi, Mô hình hoá toán học các dị thường đị hoá thứ sinh (lấy ví dụ quặng vàng vùng Đồi Bù, Lương Sơn, Hoà Bình), *Đị chất*, Cục Địa chất và Khoáng sản Việt Nam, Hà Nội, Loạt A, số 256 (2000 trang 28 - 38.

5. Belonhin M. D, Galubeva V. A, Skulop H. T, *Phân tích nhân tố trong Địa chá* NXB Nedra, Leningrat 1982, 184 trang (tiếng Nga).

6. Verkhovxkaia L. A, Cơrakina E. P., *Mô hình hoá toán học trường địa hoá vì mục đích tìm kiếm*, NXB Nedra, Leningrat 1981, 184 trang (tiếng Nga).

MATHEMATICAL MODELS OF THE ASSEMBLAGES OF ORE-INDICATING ELEMENTS IN THE SECONDARY GEOCHEMICAL FIELD OF NORTHWEST VIET NAM

Dang Mai

Department of Geology, College of Science, VNU

By application of cluster analysis according to Euclidean, chi-square, cosin Chebychev and Pearson distances, models of the assemblages of ore indicatin elements have been determined in the secondary geochemical field of Northwes Vietnam.The data on a set of metallometric samples, comprising more than 600 approximate quantitative spectral analyses have been processed.

Two main conclusions can be made on the basis of the obtained results:

1) Four assemblages of ore-indicating elements exist in the study area. They are Pb - La - Ce - Ba, Bi - Cu - W, Nb - Zr - Sn - Mo and Cr - Ni - Ti - Co - Li.

2) Pearson measure with method is suitable for the classification of the assemblages of ore - indicating elements in the secondary geochemical field.