LÝ THUYẾT NHẬN DẠNG
VÀ KHẢ NĂNG ÁP DỤNG NÓ VÀO DỰ BÁO KHÍ TƯỢNG THUỶ VĂN

Phạm Ngọc Hiện
Cục Dự báo KTTV

Tóm tắt : Bài này nhằm chủ yếu giới thiệu sơ lược về lý thuyết nhận dạng, việc vận dụng nó vào dự báo khí tượng thủy văn, đồng thời cũng công bố một kết quả vận dụng bước đầu của tác giả và đ/c Trịnh Văn Thư trong dự báo qũy đạo bão 48 giờ.

Mở đầu :

Lý thuyết nhận dạng, một lĩnh vực hiện đại bậc nhất của lý thuyết Xibecnê tic ra đời vào khoảng đầu những năm 50 của thế kỷ này dưới ảnh hưởng của các công trình của N.Vine, J.Ph. Nâyman, A.M. Turinh và U. Ros Etsobi ... và một số nhà Xibecnêtic khác. Chính N. Vine đã chỉ ra mối quan hệ chặt chẽ giữa lĩnh vực mày với các lĩnh vực kế cận của Xibecnêtic.

Dựa hẳn vào [3, 6, 10] chúng tôi muốn giới thiệu sơ lược về lý thuyết nhận dạng để các bạn làm quen với nó, và trong một chừng mực nhất định có thể tìm hiểu dễ dàng hơn một số bài báo vận dụng LTND vào dự báo KTTV.

A.- Một số kiến thức cơ bản về nhận dạng

Dạng là gì ? quả tình đến nay khái niệm này vẫn chưa có định nghĩa chính tắc. Chúng ta xem nó như một khái niệm cơ bản tiên đinh. Trên quan điểm lý thuyết nhận dạng chúng ta tạm thời hiểu dạng như một khách thể nào đó mà sau khi đã tiếp xúc (đã học nó) một số lần thì ta có thể nhận ra được nó.

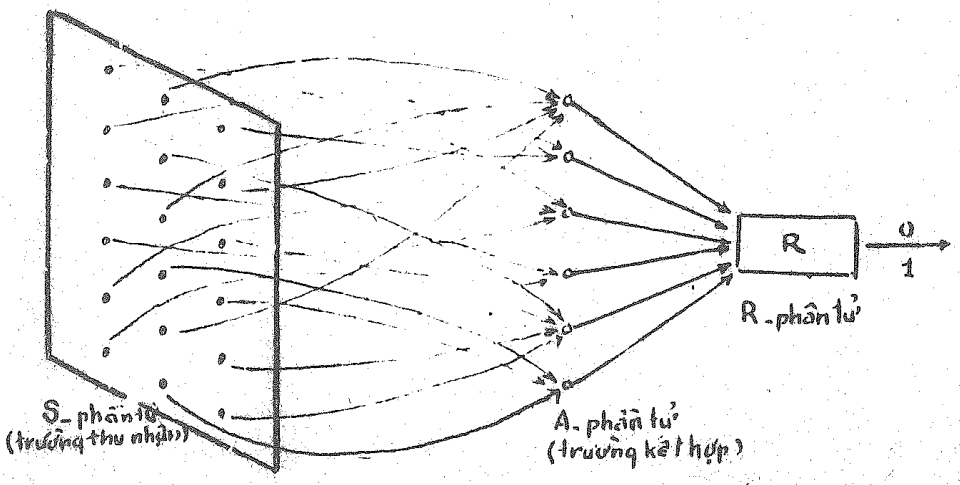
Trên nghĩa đó mọi tứ giác lồi có "dạng" tứ giác, còn tập thể những người làm "quan" chẳng phải là một dạng, mặc dầu từ bao đời nay cha ông ta vẫn nói"Làm quan có dạng làm dáng có hình" (ở đây ta loại bỏ tính liều lĩnh của các nhà bói toán). Cái thuộc tính, cái bản chất cố hữu tồn tại khách quan bên trong khái niệm dạng thì vốn có, nhưng với quan điểm ở đây rõ ràng rằng phạm trù dạng sẽ ngày một mở rộng do khả năng nhận thức thế giới khách quan của con người ngày một cao, ngày một sâu sắc.

Trước khi các máy nhận dạng ra đời và ngay cả sau khi nó đã có , não bộ con người là một "máy nhận dạng" tuyệt vời. Chính vì vậy, các thiết bị nhận dạng và các lý thuyết về nó đều mô phỏng quá trình nhận thức của hệ thần kinh. Để thấy rõ các mô phỏng đó chúng ta hãy bắt đầu từ những định nghĩa và những khái niệm cơ bản.

Ta hiểu một hệ thống, một máy là một mô hình thực hiện một ánh xạ đơn trị một tập hợp (tập hợp các xung vào hay là các tín hiệu) vào (*) một tập hợp khác(tập hợp các xung ra). Dương nhiên tín hiệu vào hoặc ra có thể là một đại lượng vật lý nào đó. Tín hiệu vào được gọi là Ảnh, tín hiệu ra gọi là trả lời của hệ thống . Hệ thống nhận dạng (hay thiết bị phân lớp) là một hệ thống có khả năng phân lớp một tập hợp ảnh phù hợp với một phân lớp tiên nghiệm nào đó. Trong khi chờ đợi một định nghĩa chính tắc hơn chúng ta hiểu phân lớp là sự phân hoạch một phần, hay một bộ phận hoặc toàn bộ các ảnh thành các tập con không giao nhau mà ta gọi là lớp các ảnh. Như vậy một hệ thống sẽ là hệ thống nhận dạng nếu nó ứng mỗi ảnh với một lớp tiên nghiệm của nó. Một hệ thống bất kỳ có khả năng trở thành hệ thống nhận dạngsau khi đưa vào nó (cho nó học, hay luyện nó) một lượng thông tin nào đó về ảnh hay một bộ phận của ảnh và các lớp ảnh được gọi là hệ thống luyện nhận dạng. Chúng ta hãy tìm hiểu một máy nhận dạng đơn giản nhất,đó là perceptron của F.Rosenblat (Mỹ), sau này nó được Gambo (Ý) cải tiến.

A.- Perceptron.

Toàn bộ Perceptron có 512 phần tử. Mỗi A - phần tử được nối ngẫu nhiên với 10S - phần tử của trường thu nhận, tức là mỗi A - phần tử có 10 cửa vào.Một nửa cửa vào ở chế độ ức chế với hệ số khuếch đại -1, một nửa ở chế độ hưng phấn với hệ số khuếch đại 1. Ở Perceptron của F.Rosenblat ngưỡng vận hành được chọn là $\lambda_0 = 0$. Tín hiệu ở cửa ra của mọi A - phần tử được đặt lên cửa vào của phần tử phản ứng gọi là R - phần tử.. Perceptron của Rosenblat chỉ nhận dạng hai ảnh



Hình 1. Perceptron của F. Rosenblat.

Thay cho việc giới thiệu quá trình vận hành của perceptron chúng tôi xin dẫn ra một thí dụ giản đơn và gần gũi hơn đối với độc giả. Thí dụ : - một máy phân biệt địch ta (máy hỏi) ở các chỉ huy sở phòng không chúng ta là một máy thu vô tuyến . Mỗi máy bay trong chiến tranh khi vận hành đều có mang theo một máy phát một tần.

(*) Ở đây hiểu theo nghĩa thông thường chứ không phân biệt ánh xạ "lên" và ánh xạ "vào" với nghĩa ánh xạ trần và ánh xạ chứng.

số nào đó. Khi một mục tiêu xuất hiện trên màn huỳnh quang người lính tiêu đồ bật máy hỏi nghĩa là bật máy thu một tần số quy định cơ mật trước, nếu máy bay là của mình thì máy thu có tín hiệu vào và phía dưới mục tiêu trên màn huỳnh quang xuất hiện một chấm sáng. Như vậy ta có một máy nhận dạng khi đã cố định tần số thu chỉ nhận dạng hai khái niệm : "Ta" và "Không phải ta".

Ta đã làm quen với một máy luyện và phần nào mường tượng ra thông qua ví dụ trên. Cần phải biết rằng một máy luyện nhận dạng chỉ có thể vận hành được với ba điều kiện chủ yếu sau đây :

Điều kiện 1 . Máy phải có tính chất ngoại suy, tính chất này có được dựatrên tính compact của tập ảnh, nói rõ hơn là tính compact của tập mô tả ảnh trong không gian mô tả. Nhờ tính chất này máy có thể luyện khi tập ảnh không đầy đủ và người ta có thể luyện máy trong tình huống một tập ảnh nghèo nàn về dung lượng.

Điều kiện 2. Hệ thống phải có khả năng thay đổi cấu trúc của nó khi nó tiếp nhận những thông tin khác nhau và do đó thực hiện các quá trình luyện khác nhau để rồi tạo ra những máy nhận dạng khác nhau. Tóm lại là hệ thống phải có tính chất vạn năng. Quá trình đưa thông tin về ảnh vào hệ thống và sự thay đổi cấu trúc hệ thống phù hợp với nó gọi là quá trình luyện hệ thống. Những ảnh dùng để luyện hợp thành liệt luyện.

Điều kiện 3 . Do tính vạn năng của hệ thống máy phải có khả năng phân lớp có quy tắc. Các ảnh sinh ra do kết quả luyện chứ không đòi hỏi đưa toàn bộ thông tin tiên nghiệm về ảnh vào máy, thực ra việc này khó làm nổi.

Nói tóm lại, hệ thống luyện nhận dạng có thể xem như một hệ thống phụ thuộc một tham số nào đó, tham số đó xác định cấu trúc của hệ thống và khi cố định một giá trị cụ thể của tham số ta có một hệ thống nhận dạng. Rất nhiều khi chúng ta gặp những bài toán mà các thuật toán luyện đối với nó đòi hỏi quá nhiều thời gian , một bài toán nhận dạng như vậy được gọi là bài toán khó. Chúng tôi cho rằng ở đây để giải nó phải quán triệt quan điểm gợi mở để tạo ra thuật toán giải nó với một độ chính xác thích hợp.

Ngoài chế độ luyện của hệ thống nhận dạng còn có chế độ thi. Đó là chế độ kiểm tra xem tần suất trả lời sai của hệ thống nhận dạng với những liệt ảnh (mà ta gọi là liệt thi) trước kia không có mặt trong liệt ảnh luyện.

B.- Những vấn đề cần biết khi vận dụng lý thuyết nhận dạng vào thực tiễn dự báo

1. Hiểu ảnh, mô tả ảnh và phương pháp mô tả ảnh trong lý thuyết nhận dạngnhư thế nào.

Ảnh đóng vai trò một tín hiệu trừu tượng đặt lên thiết bị vào của hệ thống. Các tín hiệu này được thu bởi một nhóm hữu hạn các dấu hiệu. Dấu hiệu ở đây hiểu như một hàm thực xác định trên tập hợp ảnh. Nhóm các giá trị của tất cả các dấuhiệu ứng với một ảnh gọi là một mô tả ảnh. Mô tả ảnh được xem như một điểm trong không gian dấu hiệu R^n. Với chúng ta, thường chỉ làm ứng dụng chúng ta đồng nhất ảnh với mô tả ảnh và bài toán dự báo của chúng ta thường bắt đầu từ đây. Do vậy ta cũng hiểu lớp ảnh và lớp các mô tả ảnh là một. Cần phải chú ý rằng trong thực tế do cách chọn

tham số (các dấu hiệu) thường dẫn đến chỗ làmặc dầu 2 lớp ảnh $\Omega 1$ và $\Omega 2$ không giao nhau, nhưng 2 lớp mô tả ảnh tương ứng của chúng là X1 và X2 lại có thể giao nhau.

Có nhiều cách mô tả ảnh nhưng nói chung người ta sử dụng hai phương pháp chủ yếu sau đây

Phương pháp 1 .- Phương pháp này dùng chính ngay những đại lương vật lý của ảnh làm môtả ảnh, phương pháp này giữ được nguyên vẹn thông tin nhưng nhiều k hi trong kỹ thuật cũng như trong tính toán khó thực hiên. Thí dụ . Lớp nước ở 100^oC và lớp nước ở 99,99 .... 9^oC thật là khó phân biệt,nhưng hiên tượng sôi và không sôi thì khác nhau rất xa.

Phương pháp 2 .- Phương pháp nhi nguyên hóa. Phương pháp này có thể làm mất mát một lượng thông tin nào đó nhưng nó có thể mô tả dễ dàng và thuận lợi những đặc trưng vật lý (bằng cách phân cấp cho dấu hiêu) và những đặc trưng phi vật lý (như đã làm ở t.ên).

Chẳng hạn nếu ta muốn có một mô tả ảnh cho tình thế synop : gió NE, v = 7m/sec, $\Delta T_{24} = 5^oC$; R = 75% thì phương pháp 1 cho ta mô tả ảnh (135, 7, 7, 75) tức là cho trong IR^4. Ở phương pháp 2 nếu ta thừa nhận gió có 4 phương chính và 4 phương phụ (E, W, S, N SE, NE, NW, SW), tốc độ có 3 cấp v < 4m/sec, 4 m/sec $\leq$ v $\leq$10m/sec, v > 10 m/sec; ΔT_{24} với 3 cấp $\Delta T_{24} < 1^oC$, $1^oC \leq \Delta T_{24} \leq 5^oC$, $\Delta T_{24} > 5^oC$; R với các cấp R<60%; 60% $\leq$ R $\leq$ 90%, R>90% thì mô tả ảnh sẽ là :

$$(0,0,0,0,0,0,1,0,0,0,1,0,0,1,0,0,1,0) \in IR^{17}$$

Cùng với 2 phương pháp trên là các biến tướng của nó độc giả có thể tìm thấy ở bất kỳ cuốn sách nào về nhận dạng.

Trong bài toán dự báo thời tiết thông thường người ta nhận dạng các tình thế synốp, đương nhiên không phai người ta chỉ bó hẹp các dấu hiêu là các số liệu synop.

Rõ ràng rằng để mô tả ảnh nói chung, và mô tả ảnh cho một tình thế synop không khó nhưng việc chọn các dấu hiệu phản ánh bản chất nó thì thật là khôn cùng.

Sau khi đã chon được một số tham số, chúng ta cần phải hiểu rằng trong các tham số đó có những tham số phan ánh bản chất của ảnh, có những tham số có khichẳng có một ý nghĩa nào. Vì vậy cũng như những bài toán thống kê khác chúng ta phải lọc tham số.

Quy tắc quyết định.

Bài toán nhận dang nói chung đưa về bài toán tìm quy tắc quyết định. Nó được phát triên theo 3 hướng [2] khác nhau. Đó là :

i) Hướng thứ nhất xuất hiện các công trình đặt giả thuyết cho h à m quyết định : tuyến tính, tuyến tính từng khúc, phi tuyến.

ii) Hướng thứ hai xuất phát điểm là các phán đoán về tín hiêu, mô hình tín hiêu.

iii) Hướng cuối cùng là ng theo quan điểm gợi mở (heuristic) thể h i ệ n nguyện vọng tìm một thuật toán giải cho kỳ được bài toán nhận dạng hay

luyện nhận dạng dựa trên những dự kiện trực giác (nói chung không chặt chẽ về mặt toán học). Về mặt kỹ thuật [illegible] hiện những thiết bị chuyên dụng; cả về mặt kỹ thuật cũng như về mặt toán học kết quả của hướng này phụ thuộc rất nhiều vào tài năng và linh cảm của người tạo ra nó.

Trên quan điểm nguyên lý mà nói lý thuyết nhận dạng có 2 hướng phát triển chủ yếu [9] :

i) Hướng thứ nhất, nghiên cứu các phương pháp nhận dạng có bản chất con người và mọi cơ thể sống khác.

ii) Hướng thứ 2. Phát triển lý thuyết và các phương pháp xây dựng các thiết bị dùng để giải các bài toán riêng trong những lĩnh vực thực nghiệm riêng.

Như vậy là hướng này trùng với hướng thứ 3 ở trên. Chúng tôi cho rằng những người làm dự báo nên và đã đứng trước hướng thứ 2 đầy triển vọng này. Nhiệm vụ của chúng ta là dùng máy tính điện tử tạo ra hệ thống nhận dạng các tình thế synop [4, 5, 8] .Nhưng mọi người khi bắt tay vào lý thuyết nhận dạng cần phải lý giải thậtngã ngũ trong lòng câu hỏi "Vậy lý thuyết nhận dạng là vấn đề khoa học hay chỉ là gợi mở ! "Chúng tôi xin giành cho các bạn câu trả lời sau khi tìm hiểu[7] của V.A.Ko - valepski.

3. Hàm phân lớp .

Gọi $\Omega_1, \Omega_2, \ldots, \Omega_l$ là các tập con không giao nhau của tập trừu tượng Ω các ω . Khi đó các Ω_i (i = 1, 2, ...l) được gọi là lớp các ảnh.

Mọi hệ hàm thực $\{S_i(\omega)\}$; i = 1, 2, ... l thỏa mãn các tính chất.

$$S_K(\omega) > \max_{i \neq K} Si(\omega) \quad \text{với mọi} \quad \omega \in \Omega_K \quad (*)$$

được gọi là hàm phân lớp ứng với các tập $\Omega_1, \Omega_2, \ldots, \Omega_l$. Một trường hợp cực kỳ quan trọng mà mọi trường hợp dẫn về nó là trường hợp l = 2,khi đó hàm

$$S(\omega) = S_1(\omega) - S_2(\omega)$$

chính là hàm phân lớp 2 tập trên. Ở đây $\forall \omega \in \Omega_1$ ta có $S(\omega) > 0$; $\forall \omega \in \Omega_2$ ta có $S(\omega) < 0$. Tức là Ω_1 và Ω_2 được tách ra theo dấu của $S(\omega)$.

Vậy xây dựng trước một quy tắc ứng mỗi ảnh với một tập con không giao nhau của hệ $\{\Omega_i\}$ (i = 1,2, ..., l) được gọi là một phân lớp tiên nghiệm. Như vậy một phân lớp tiên nghiệm sẽ được xác định hoàn toàn bởi cách xây dựng các hàm $\{s_i(\omega)\}$ (i = 1, 2, ..., l).

Ở Perceptron nếu trên Ω có cho hệ hàm thực $x^1(\omega), \ldots, x^m(\omega)$. Tức là cho ánh xạ

$$X(x^1, \ldots, x^m) : \Omega \Rightarrow \mathbb{R}^m$$

thì

$$X : \Omega_i \rightarrow X_i$$

Khi đó Ω đóng vai trò các S - phần tử, còn $\mathbb{R}^m$ đóng vai trò là không gian các dấu hiệu sẽ đóng vai trò các tín hiệu vào của A - phần tử và $(x^1(\omega), \ldots, x^m(\omega))$ gọi là mô tả ảnh của ω.

B.- Nhận dạng và dự báo KTTV . Một kết quả thử nghiệm (*)

1. Tình hình chung .-

Nói chung, lý thuyết nhận dạng là một vấn đề lớn, một ngành mới mẻ, nó đáp ứng được nhiều yêu cầu của nhiều ngành khoa học hiện đại. Trong lĩnh vực dự báo khí tượng thủy văn phần áp dụng chủ yếu là phương pháp luyện nhận dạng. Vấn đề luyện nhận dạng có thể xem như có hai hướng chủ yếu :

i) Hướng thứ nhất thường gán cho tập ảnh (tập mô tả ảnh) một phân bố xác suất nào đó - chủ yếu là phân bố chuẩn - và thủ tục luyện như vậy gọi là thủ tục luyện nhận dạng có tham số (tham số phân bố xác suất).

ii) Hướng thứ hai là hướng không gán cho tập ảnh một phân bố nào cả, thủ tục luyện trong trường hợp này gọi là thủ tục luyện nhận dạng phi tham số.

Trong những bài toán dự báo khí tượng thủy văn thông thường chúng ta có một số số liệu lịch sử đủ lớn. Tình hình ấy cho phép chúng ta gán cho dãy số liệu một phân bố nào đó và kiểm chứng được giả thiết Thống kê cho nó. Đ/c Nguyễn Viết Phong (trong dự báo gió mùa đông bắc 24 h), Ngô Ngọc Thạch (trong dự báo gió mùa đông bắc 24h và 48h), và Nguyễn Văn Tuyên (trong dự báo giáng thủy) đã đạt được một số kết quả nhất định theo hướng này(**).

Khi số liệu lịch sử quá ít ỏi (tình trạng này sảy ra đối với những hiện tượng thời tiết nguy hiểm và đặc biệt nguy hiểm) thì chúng ta không thể gán cho tập số liệu đó một phân bố xác suất nào, vì lẽ việc xác định các tham số thống kê đối với nó sẽ không chính xác. Trong tình trạng ấy người ta buộc lòng phải tiến hành luyện nhận dạng phi tham số. Ở Liên xô A.N. Bagrop, V.N. Vapnhic, A.T. Snhicovưi đã bàn rất nhiều về dự báo hiện tượng thời tiết nguy hiểm [4]. G.M. Vinograxop dùng lý thuyết luyện nhận dạng phi tham số để dự báo nhiệt độ tối thấp nhờ phương pháp ảnh suy rộng [5], T.B.Rostcova và A.E.Oocdanôvic còn dùng luyện nhận dạng vào việc nghiên cứu chuyển động của xoáy thuận nhiệt đới [8]. Nhìn chung những công trình theo hướng này chưa thật rộng và các kết quả mà nó đạt được theo nhận thức của cá nhân khó mà vượt được những kết quả của hướng thứ nhất.

Áp dụng vào thực tế, chúng tôi chọn phương pháp luyện nhận dạng bằng phương pháp ảnh suy rộng làm công cụ thử nghiệm lập một mô hình dự báo khuynh hướng di chuyển của bão trong 48h tới.

(*) Kết quả này do Phạm Ngọc Hiên và Trịnh Văn Thư cộng tác.

(**) Chú ý rằng lý thuyết hàm phân lớp ra đời trước khi lý thuyết nhận dạng hình thành.

2. Ứng dụng phương pháp luyện nhận dạng nhờ ảnh suy rộng dự báo khuynh hướng di chuyển của bão trong 48h tới

a) Sơ lược về phương pháp ảnh suy rộng

Giả sử ta có hai tập hữu hạn các vectơ mô tả ảnh X_a và $\bar{X}_b$ trong không gian $\mathbb{R}^K$

$$X_a\ (x_1, x_2, \dots, x_a) \qquad (1)$$

$$\bar{X}_b\ (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_b) \qquad (2)$$

Để lập một siêu phẳng phân lập 2 tập trên tức là lập hàm phân lớp chúng ta thiết lập hai bài toán sau đây :

Bài toán 1 . (Tìm ảnh suy rộng).

Tìm một vectơ $\Psi \in \mathbb{R}^K$ sao cho :

$$\left.\begin{array}{ll} (x_i, \Psi) \geq 1 & \{\forall x_i \in X_a\} \\ (\bar{x}_j, \Psi) \leq K & \{\forall \bar{x}_j \in X_b\};\ K < 1 \end{array}\right\} \qquad (3)$$

Như chúng ta biết số vectơ Ψ như vậy là vô số và nó phụ thuộc vào K, cho nên đúng ra phải viết là $\Psi(K)$. Để chọn được $\Psi(k)$ đảm bảo (3) không có sai lầm thực ra phải đòi hỏi điều kiện (A. b. 4i) (*) ở phần trước (Trong thực tế của chúng ta điều kiện này nói chung là không thỏa mãn). Bài toán của ta là tìm một vectơ cực tiểu về module trong họ $\Psi(K)$ thỏa mãn (3). Vectơ đó được gọi là ảnh suy rộng của tập X_a đối với tập $\bar{X}_b$ (khái niệm Ảnh suy rộng nảy sinh từ bản chất nó là trung bình theo nghĩa minmax (Ψ, x) [3]). Như vậy bài toán 1 có thể được phát biểu lại như sau :

Tìm $\Psi \in \mathbb{R}^K$ sao cho

$$\|\Psi\|^2 = \min_K \|\Psi(K)\|^2 \qquad (4)$$

thỏa mãn các điều kiện (3).

Đó là một bài toán quy hoạch toàn phương, sự tồn tại và duy nhất nghiệm của nó được đảm bảo bởi định lý Kuhn - Tukker. Khi đó siêu phẳng phân lập sẽ là

$$(x, \Psi) = \frac{1+K}{2} \qquad (5)$$

Thông thường người ta không trực tiếp giải bài toán 1 mà giải bài toán đối ngẫu với nó.

Bài toán 2 (bài toán đối ngẫu của bài toán trên).

Trong không gian các tham số $E_{\alpha\beta}$ tìm cực đại hàm

$$W(\alpha, \beta) = \sum_{i=1}^{a} \alpha_i - K \sum_{j=1}^{b} \beta_j - \frac{1}{2} (\Psi, \Psi) \qquad (6)$$

(*) A. Phần A, mục b, điểm thứ 4 và điều kiện i.

ở đây

$$\Psi = \sum_{i=1}^{a} \alpha_i \chi_i - \sum_{j=1}^{b} \beta_j \bar{x}_j \qquad (7)$$

với $\alpha_i \geqslant 0$; $\beta_j \geqslant 0$.

Trong [1, 6] V.N. Vapnhic và A.Ia Trevonhenkis đã chứng minh rằng hai bài toán này là tương đương.

Trong thực hành bài toán này được giải bằng phương pháp Gradunt - liên hợp [6] để tìm nghiệm gần đúng của nó, tức là chỉ tìm α_i và β_j (tức là tìm Ψ theo (7)) thỏa mãn các điều kiện :

$$\frac{\partial W(\alpha, \beta)}{\partial \alpha_i} < \frac{1-k}{2}$$

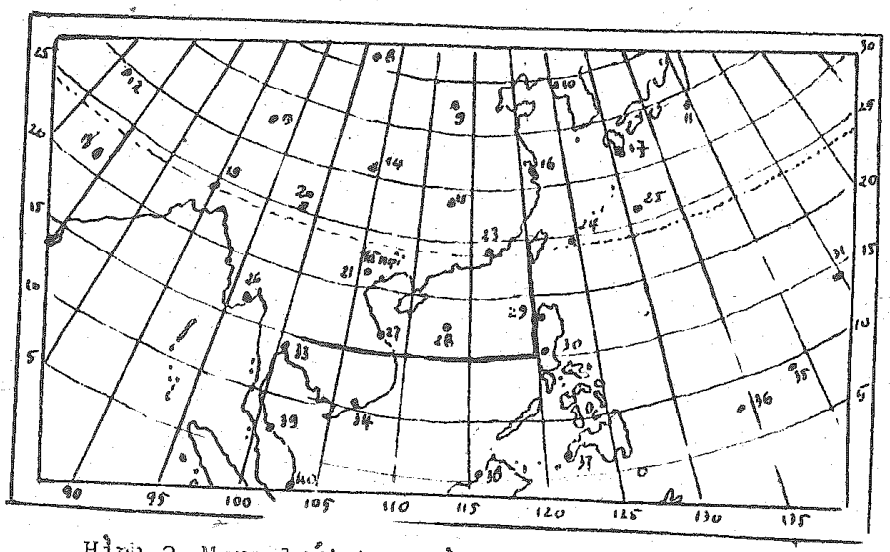
$$\frac{\partial W(\alpha, \beta)}{\partial \beta_j} \leqslant \frac{1-K}{2}$$

b) Ứng dụng bài toán vào dự báo đường đi của bão.

1. Phân lớp tiên nghiệm. Chúng tôi chọn lớp X_a là lớp các tình thế synôp có bão di chuyên 48giờ tới thấp hơn, tức là lệch nam) so với di chuyên quá khư của nó; $\bar{X}_b$ là lớp tình thế synôp có bão di chuyên cao hơn (tức là lệch bắc) so với di chuyên quá khứ của nó. Vì chỉ trình bày vấn đề này như một thí dụ chúng tôi xin bỏ qua việc lý giải các ý nghĩa vật lý, ý nghĩa dự báo ... của từng bước.

2. Các tham số chọn ở đây là :

$\bar{x}$ - 7 trị số đặc trưng cho Laplacient của trường độ cao H trung bình của các mặt đẳng áp 700 mb và 500 mb (tính sai khác bằng số ΔS) dựa vào mạng lưới điểm (*) của hình 1 được tính như sau :



Hình 2. Mạng lưới trạm và miền dự báo.

(*) Trị những điểm này có quan trắc cao không.

$$X_1 = \frac{1}{5}(H_9 + H_{14} + H_{16} + H_{23} + H_{21}) - H_{15}$$

$$X_2 = \frac{1}{5}(H_{14} + H_{20} + H_{22} + H_{27} + H_{29} - H_{21}$$

$$X_3 = \frac{1}{5}(H_{15} + H_{16} + H_{22} + H_{24} + H_{29}) - H_{23}$$

$$X_4 = \frac{1}{5}(H_{16} + H_{17} + H_{23} + H_{25} + H_{29}) - H_{24}$$

$$X_5 = \frac{1}{5}(H_{21} + H_{26} + H_{27} + H_{34} + H_{39}) - H_{33}$$

$$X_6 = \frac{1}{5}(H_{21} + H_{22} + H_{28} + H_{33} + H_{34}) - H_{27}$$

$$X_7 = \frac{1}{5}(H_{27} + H_{33} + H_{39} + H_{40} + H_{38}) - H_{34}$$

(Vì chọn mạng lưới chung của đề tài bão cho nen chúng tôi dùng hình 2 mặc dầu có một số điểm trên hình vẽ không dùng tới).

* - Một tọa độ vĩ tuyến của vị trí tâm bão : X_8 tại thời điểm xuất phát.

* - Tốc độ di chuyên của bão 24 giờ trước : X_9.

Chúng tôi chọn 136 trường hợp bão từ 1960 - 1974 trên khu vực biển Đôngthuộc tây 120° kinh tây và bắc 15° vĩ bắc. Chúng tôi phân thành 86 trường hợp dùng làm liệt luyên và 50 trường hợp làm liệt thi.

3. Mô tả ảnh . Ảnh với 9 dấu hiệu trên được nhị nguyên hóa thành 4 cấp

$$\left.\begin{array}{l} X_i \leqslant A_i \\ A_i < X_i \leqslant (AB)_i \\ (AB)_i < X_i \leqslant B_i \\ \text{và} \quad X_i > B_i \end{array}\right\}$$

trong đó : $A_i = \min_i (M x_i, M\bar{x}_i)$

$B_i = \max_i (Mx_i , M\bar{x}_i)$

$$(AB_i) = \frac{A_i + B_i}{2}$$

với Mx_i và $M\bar{x}_i$ là kỳ vọng toán của nhân tố thứ i của 2 lớp X_a và $\bar{X}_b$. Những trị này được cho trong bảng 1. (xem bảng 1 ở trang sau).

Khi trị của dấu hiệu rơi vào cấp nào ta cho tọa độ đó mã 1, còn cấp còn lại nhận mã 0. Như vậy ảnh nhị nguyên hóa được mô tả trong không gian IR^{36}. Một thí dụ tình huống 0000GMT ngày 3/VII/1981 của cơn bão Kelly được cho trong bảng 2.Toàn bộ quá trình này được xử lý trên MTĐT Minsk-32

4. Quá trình luyện nhận dạng.

Ở đây quá trình luyện nhận dạng chính là quá trình giải bài toán liên hợp ở trên với thuật toán có bộ nhớ đầy đủ [6]. Siêu phẳng phân lập tìm thấy trongkhông

gian R^{36} ứng với K = 0,2 là

$$\sum_{i=1}^{36} Yi \cdot \Psi i - 0,6 = 0$$

Do đó hàm phân lớp là

$$\omega = \sum_{i=1}^{36} Yi\, \Psi i - 0,6.$$

5. Dự báo

- Nếu $\omega > 0 \Rightarrow \sum Yi \cdot \Psi i - \frac{1+K}{2} > 0$

$\Rightarrow (\underline{r}, \Psi) > \frac{1+K}{2} \Rightarrow$ Tình huống dự báo thuộc lớp thứ 1 là tình huống trong 48h tới bão sẽ di chuyển thấp hơn qũy đạo của nó ở 24h trước.

- Ngược lại $\omega \leqslant 0$ là tình huống dự báo thuộc lớp thứ 2 là tình huống trong 48h tới bão di chuyển cao hơn qũy đạo của nó ở 24h trước. Trong thí dụ về bão Kelly, ta có :

$$= - 0,0638 + 0,2576 + 0,1103 + 0,1942 + 0,2356 + 0,2547 +$$

$$+ 0,3160 + 0,0724 + 0,2813 - 0,6 = 1,0683 > 0$$

Dự báo của chúng tôi là trong 48h bão lệch nam. So với thực tế dự báo của chúng tôi là đúng. (xem bảng 2 ở trang sau).

6. Đánh giá sơ bộ .

Mặc dầu dự báo sự di chuyển của bão 48h (dù chỉ là khuynh hướng) là một việc khó khăn, kết quả kiểm lại trên liệt thi (50 trường hợp) đối với lớp 1, lớp các cơn bão lệch nam là 81% còn đối với lớp lệch bắc là 64% và kết quả chung là 71 % ; cần nói rõ hơn là đánh giá này làm máy móc theo lý thuyết trên MTĐT chứ không sử dụng sai số cho phép về hướng.

Chúng tôi chưa khẳng định rõ mặt khoa học và thực tiễn của kết quả này. Nhưng chúng tôi tin rằng trong khi chúng ta chưa có MTĐT trong tay thì việc tạo ra một sơ đồ mà khi sử dụng nó nhẹ nhàng như đã thấy là một việc nên làm.

Tài liệu tham khảo

1 V.N. Vapnhic - Khôi phục các quan hệ phụ thuộc theo các số liệu t h ự c nghiệm. M. "Khoa học". 1979.

2 V.A. Kovalepsky. Phương pháp quyết định tối ưu trong nhận dạng hình ảnh , M."Khoa học" 1976.

3 V.N. Vapnhic và A.Ia. Trevonhenkis . Lý thuyết nhận dạng - M - "Khoa học" 1974.

4 A.N. Bagrop, V.N. Vapnhic, A.I. Snhitkovsky. Về dự báo thời tiết n g u y hiểm "KTTV" LX.N.3.1974.

5 G.M.Vinogradop. Thử nghiệm mô hình dự báo ngắn hạn nhiệt khí tối thấp nhờ phương pháp ảnh suy rộng "KTTV". LX. N2. 1968.

Bảng 2. Cách nhị nguyên hóa ảnh trong IR^{36} thành phần ảnh suy rộng trong $I^{p\,36}$ và một tình huống mô tả ảnh ví dụ của [illegible]n bao Kelly (0000GMT-3/VII/1981).

X	Y	Ngưỡng Y_i	Ảnh suy rộng	X (*)	Y*
X_1	y_1	$X_1 \leqslant -1,916$	-0,0638	-2,4	1
	y_2	$-1,016 < X_1 \leqslant -1,761$	-0,7744		0
	y_3	$-1,761 < X_1 \leqslant -1,60$	0,0000		0
	y_4	$X_1 > -1,606$	0,3611		0
X_2	y_5	$X_2 \leqslant -0,917$	0,2576	-3,4	0
	y_6	$-0,917 < X_2 \leqslant -0,736$	0,0390		0
	y_7	$-0,736 < X_2 \leqslant -0,518$	0,0562		0
	y_8	$X_2 > -0,548$	-0,1331		0
X_3	y_9	$X_3 \leqslant -0,600$	0,1103	-1,6	1
	y_{10}	$-0,600 < X_3 \leqslant -0,574$	0,0000		0
	y_{11}	$-0,574 < X_3 \leqslant -0,548$	0,0000		0
	y_{12}	$X_3 > -0,548$	0,1094		0
X_4	y_{13}	$X_4 \leqslant -1,511$	0,1942	-1,9	1
	y_{14}	$-1,511 < X_4 \leqslant -1,466$	0,0000		0
	y_{15}	$-1,466 < X_4 \leqslant -1,420$	0,0000		0
	y_{16}	$X_4 > -1,420$	0,0255		0
X_5	y_{17}	$X_5 \leqslant -0,796$	0,1473	1,6	0
	y_{18}	$-0,796 < X_5 \leqslant -0,542$	-0,0449		0
	y_{19}	$-0,542 < X_5 \leqslant -0,289$	0,0143		0
	y_{20}	$X_5 > -0,289$	0,2356		1
X_6	y_{21}	$X_6 \leqslant 0,208$	0,2647	-0,2	1
	y_{22}	$0,208 < X_6 \leqslant 0,299$	0,0000		0
	y_{23}	$0,299 < X_6 \leqslant 0,372$	0,0000		0
	y_{24}	$X_6 > 0,372$	-0,0449		0
X_7	y_{25}	$X_7 \leqslant 0,480$	-0,1062	1,2	0
	y_{26}	$0,480 < X_7 \leqslant 0,529$	0,0000		0
	y_{27}	$0,529 < X_7 \leqslant 0,572$	0,0000		0
	y_{28}	$X_7 > 0,572$	0,3160		1
X_8	y_{29}	$X_8 \leqslant 17,476$	0,0724	16,5	1
	y_{30}	$17,476 < X_8 \leqslant 17,516$	0,0312		0
	y_{31}	$17,516 < X_8 \leqslant 17,556$	0,0000		0
	y_{32}	$X_8 > 17,556$	0,1161		0
X_9	y_{33}	$X_9 \leqslant 2,992$	-0,0658	4,6	0
	y_{34}	$2,992 < X_9 \leqslant 3,132$	-0,0257		0
	y_{35}	$3,131 < X_9 \leqslant 3,272$	0,0300		0
	y_{36}	$X_9 > 3,272$	0,2813		1

(*) Dấu * chỉ trường hợp thí dụ.

(xem tiếp trang 30)