

Следващата стъпка е построяването на матрицата на разстоянията между признаците, която е в основата на определяне на групите от сходни признаци, и изчисляването на коефициентите на значимост⁵ за всеки от признаците (Плюта, 1980, с. 42). Всеки елемент от матрицата се получава по формулата:

$$c_{pq} = \sqrt{\frac{\sum_{i=1}^w (z_{pi} - z_{qi})^2}{w}}$$

където:

$p, q = 1, 2, \dots, n$; w е броят на единиците на наблюдението (държавите).

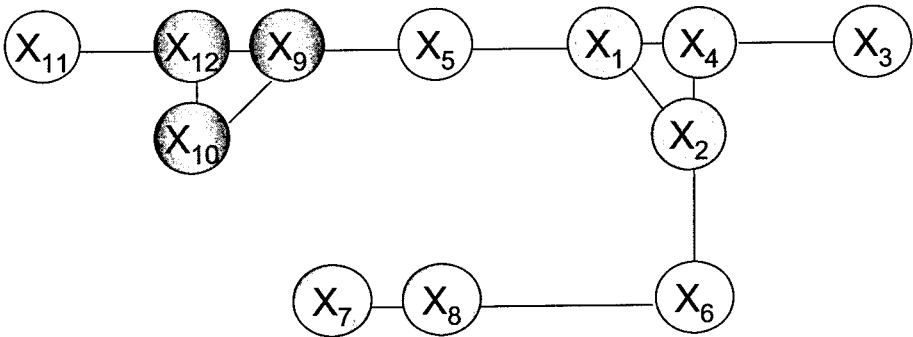
Определянето на групите със сходни признаци се основава на минималното

разстояние между тях на всеки ред от матрицата на разстоянията: $c_i = \min_p c_{ip}$.

В състава на еднородните подмножества влизат сходни елементи с близки значения на признаците. Групирането на признаците, съответстващо на оптималното подреждане, трябва да отговаря на условието за минималното разстояние между всеки два съседни признака (съседните елементи в най-малка степен се отличават един от друг).

След изпълнение на алгоритъма по построяване на матрицата на разстоянията между признаците се получава следният модел за връзка - т.нар. дендрит (пак там, с. 24).

Фиг. 1. Модел за връзка между признаците за оценка на образователното равнище на българските емигранти за 1990 г.



Процедурата по обединяването на признаците съответства в определена степен на установяването на корелационна зависимост (колинеарност) между факторните променливи в модела. Обединяването на признаците в еднородни групи трябва да отговаря на две условия - силна корелационна зависи-

мост между признаците във всяка група и слаба корелация между факторните променливи, принадлежащи към различни групи.

Всяка от получените групи признаци характеризира определен аспект на образователното равнище на емигрантите. Първата група (x_9, x_{10} и x_{12}) включва

⁵ В литературата се среща като коефициент на йерархия на признаците.