ЛАБОРАТОРНА РОБОТА № 8
РОЗРОБКА СИСТЕМ АНАЛІЗУ ДАНИХ МЕТОДАМИ НЕЧІТКОЇ КЛАСТЕРІЗАЦІЇ

Мета роботи: Освоїти методику проектування і побудови систем інтелектуального аналізу даних на основі методів нечіткої кластеризації.

8.1. Основні поняття

Кластерний аналіз – це сукупність методів, підходів і процедур, які розробляються для розв'язування проблеми формування класів – сукупностей даних, однорідних за заданими ознаками.

Кластерний аналіз (автоматична класифікація сукупності даних) займає одно з центральних місць серед методів аналізу даних і являє собою сукупність підходів та алгоритмів знаходження деякого розбиття досліджуваної сукупності об'єктів на підмножини відносно схожих між собою елементів. Такі підмножини отримали назву кластерів.

Виділення кластерів серед сукупності даних має відповідати наступним вимогам:

1. кожний кластер представляє собою сукупність об'єктів, які схожі між собою значеннями деяких властивостей або ознак;
2. сукупність всіх кластерів має бути вичерпаною, тобто всі об'єкти досліджуваної сукупності мають належить до деякого кластеру;
3. кластери мають бути взаємно-виключні; тобто, жоден з об'єктів не має належить до двох різних кластерів.

Формально, під задачею кластерного аналізу розуміється задача знаходження деякого теоретико-множинного розбиття початкової множити об'єктів на підмножини, які не перетинаються, таким чином, щоб елементи, які відносяться до однієї підмножини відрізнялися між собою в значно меншій степені, ніж об'єкти з різних підмножин.

Концептуальний зв'язок між кластер ним аналізом і теорією нечітких множин оснований на тому, що при розв'язуванні задач структуризації складних систем більшість класів, що формуються, «розмиті» за своєю природою. Тому, найбільш адекватну відповідь слід шукати не на питання «Чи належить елемент до того чи іншого класу?», а на питання «В якій степені даний елемент належить класу, що розглядається?».

Методи нечіткої кластеризації вводять до розгляду нечіткі кластери і відповідні їм функції належності, які приймають значення з інтервалу $[0,1]$.

Таким чином, задача нечіткої кластеризації полягає у тому, що необхідно знайти нечітке розбиття або нечітке покриття множини елементів сукупності, що досліджується. Задача зводиться до знаходження степенем належності елементів множини нечітким кластерам (класам).

8.2. Постановка задачі

Нехай початкова (досліджувана) сукупність даних представляє собою скінчену множину елементів $A = \{a_1, a_2, ..., a_n\}$, яке ще називається множиною об'єктів кластеризації. Вводиться також скінчена множина ознак або атрибутів об'єктів $P = \{p_1, p_2, ..., p_q\}$, кожний з яких являє собою деяку характеристику елементів множини A.

Далі, пропонується, що для всіх елементів множини об'єктів кластеризації виміряли всі ознаки множини P, і кожен елемент множини $a_i \in A$ представлений вектором $x_i = \{x_i^1, x_i^2, ..., x_i^q\}$, де $x_i^j \in R$ - дійсне значення ознаки $p_j \in P$ для об'єкту $a_i \in A$.

Взагалі, проблема кількісного вимірювання ознак кожного об'єкта з сукупності – нетривіальна і самостійна задача. Процес вимірювання ознак може бути реалізований в різних шкалах, кожна з яких характеризується допустимим перетворенням даних. В зв'язку з цим, визначають різні типи шкал:

- шкала найменувань: об'єкту ставиться у відповідність деякий символ або номер, який лише відокремлює одне значення ознаки від іншого; прикладом таких ознак є стать людини –(м, ж) або найменування міст (Київ, Житомир, Луганськ,...); допустимим відображенням множини об'єктів у множину символів є бієктивне відображення;
- порядкова шкала: разом з відповідною множиною символьних ознак об'єктів ця шкала дозволяє встановити відношення порядку відносно цієї ознаки; тоді об'єкту ставиться у відповідність деяке число, яке грає роль його оцінки в балах; допустимим відображенням є монотонне зростаюче відображення або функція між двома множинами значень ознак; приклад – оцінки на іспитах;
- інтервальна шкала: крім порядку елементів по ознакам ця шкала встановлює рівність інтервалів значень цієї ознаки; об'єкту, як правило, ставиться у відповідність число, яке дорівнює значенню цієї ознаки; допустимим перетворенням тут є довільна лінійна зростаюча функція між двома множинами значень ознак; характерною ознакою такої шкали є відсутність абсолютного нуля; приклад – температура в шкалах Цельсія;
- шкала відношень: в доповнення до рівності інтервалів додає ще рівність відношень значень ознаки, що розглядається; об'єкту ставиться у відповідність деяке число, яке дорівнює значенню цієї ознаки; допустимим відображенням є довільна лінійна зростаюча функція, яка проходить через нуль; приклад – відстань в метрах, швидкість в км/ч.

Множину ознак слід обирати таким чином, щоб всі $x_i^j \in R$ були виміряні в шкалах відношень чи інтервалів. Саме в такому випадку результати нечіткої кластеризації мають змістовну інтерпретацію, яка адекватна проблемі знаходження нечітких кластерів.

Вектори значень ознак $x_i = \{x_i^1, x_i^2, ..., x_i^q\}$ зручно представляти у вигляді матриці даних D розмірності (n*q), кожний рядок якої представляє собою значення вектору x_i .

Отже, **задача нечіткого кластерного аналізу формулюється наступним чином**: на основі даних матриці D визначити таке нечітке розбиття $R(A)=\{A_k \mid A_k \subseteq A\}$ або нечітке покриття $J(A)=\{A_k \mid A_k \subseteq A\}$ множини A на задане число нечітких кластерів $A_k(k \in \{2,...,c\})$, яке доставляє екстремум деякій цільовій функції $F(R(A))$ серед всіх нечітких розбиттів чи екстремум цільової функції $F(J(A))$ серед всіх можливих нечітких покриттів.

Для конкретизації задачі ще слід уточнити вигляд цільової функції та тип шуканих нечітких кластерів.

Одним з видів конкретизації цієї задачі є використання спеціальної функції fcm системи MATLAB, який оснований на алгоритмі розв'язування методом нечітких с-середніх.

Для уточнення вигляду цільової функції $F(J(A))$ вводяться деякі додаткові поняття. По-перше, пропонується, що шукані нечіткі кластери представляють собою нечіткі множини $A_k(k \in \{2,...,c\})$, які є нечітким покриттям початкової множини об'єктів кластеризації A, для якої має місце наступна умова:

$$\sum_{k=1}^{c} \mu_{A_k}(a_i)=1,(\forall a_i \in A), \qquad (8.1)$$

де c – загальна кількість нечітких кластерів $A_k(k \in \{2,...,c\})$, яке вважається попередньо заданим.

Далі для кожного кластеру вводяться так звані типові представники або **центри** v_k шуканих нечітких кластерів $A_k(k \in \{2,...,c\})$, які розраховуються за наступною формулою:

$$v_j^k=\frac{\sum_{i=1}^{n}(\mu_{A_k}(a_i))^m x_i^j}{\sum_{i=1}^{n}(\mu_{A_k}(a_i))^m},(\forall k \in \{2,...,c\},\forall p_j \in P), \qquad (8.2)$$

де m – деякий параметр, який має назву **експоненційна вага** і дорівнює деякому дійсному числу (m>1). Кожний з центрів кластерів є вектором $v_k=(v_k^1,v_k^2,...,v_k^q)$ в деякому q-вимірному нормованому просторі ознак, який ізоморфний R^q, якщо всі ознаки виміряні по шкалі відношень.

В якості цільової функції будемо розглядати суму квадратів зважених відхилень координат об'єктів кластеризації від центрів нечітких кластерів:

$$F(A_k, v_k^j)=\sum_{i=1}^{n}\sum_{k=1}^{c}(\mu_{A_k}(a_i))^m \sum_{j=1}^{q}(x_i^j-v_k^j)^2 . \qquad (8.3)$$

Чим більше елементів містить множина A, тим менше значення слід вибирати для m>1.

Тоді задача нечіткої кластеризації полягає у наступному: для заданої матриці даних D, кількості нечітких кластерів $c \in N, c>1$, параметра m, визначити матрицю U значень функції належності об'єктів кластерізації $a_i \in A$ нечітким кластерам $A_k(k \in \{2,...,c\})$, які доставляють мінімум цільової функції (8.3) і задовольняють обмеженням (8.1)-(8.2), а також додатковим обмеженням:

$$\sum_{i=1}^{n} \mu_{A_k}(a_i) > 0, (\forall k = \{2,..., c\})$$
$$\mu_{A_k}(a_i) \geq 0 (\forall k = \{2,..., c\}, a_i \in A) \quad (8.4).$$

Умови (4) виключають появу пустих нечітких кластерів в шуканій нечіткій кластеризації. Таким чином, мінімізація цільової функції (8.3) мінімізує відхилення всіх об'єктів кластеризації від центрів нечітких кластерів пропорційно значенням функцій належності цих об'єктів відповідним нечітким кластерам.

Ця функція не є випуклою, а тому задача кластеризації в загальному випадку відноситься до багатоекстремальних задач нелінійного програмування.

8.3. Алгоритм розв'язування задачі нечіткої кластеризації

Основні ідеї алгоритму для розв'язування задачі нечіткої кластеризації були запропоновані Дж.К.Данном у 1974р. Цей алгоритм спочатку отримав назву нечіткого алгоритму fuzzyISODATA. У 1980 році Дж.К. Беджек теоретично довів збіжність цього алгоритму, потім він же узагальнив цей алгоритм на випадок довільних нечітких множин даних і запропонував для цього алгоритму назву нечітких середніх FCM, Fuzzy-C-Means. Саме під такою назвою алгоритм реалізований в системі MATLAB.

Алгоритм FCM має ітеративний характер послідовного покращення деякого початкового нечіткого розбиття $R(A) = \{A_k \mid A_k \subseteq A\}$, яке задається користувачем або формується автоматично за деяким евристичним правилом. На кожному кроці ітерації прораховуються значення функцій належності нечітких кластерів і їх типових представників.

Алгоритм FCM завершує роботу у випадку, коли відбудеться наперед задане число ітерацій, або, коли мінімальна абсолютна різниця між значеннями функцій належності на двох послідовних ітераціях не стане менше деякого наперед заданого значення.

Формально алгоритм FCM представляється у вигляді наступних кроків:

1. попередньо необхідно задати наступні значення: кількість шуканих нечітких кластерів c, максимальну кількість ітерацій алгоритмів $s \in N$, параметр збіжності алгоритму $\varepsilon \in R_+$, а також екпоненційну вагу для цільової функції і центрів кластерів m>1. В якості початкового розбиття на першій ітерації алгоритму для матриці даних D задати деяке нечітке розбиття $R(A) = \{A_k \mid A_k \subseteq A\}$ на c непустих нечітких кластерів, які описуються сукупністю функцій належності $\mu_k(a_i), \forall k \in \{2,..., c\}, \forall a_i \in A$.
2. для поточного нечіткого розбиття $R(A) = \{A_k \mid A_k \subseteq A\}$ за формулою (8.2) розрахувати центри нечітких кластерів $v_k^j, (\forall k = \{2,..., c\}, \forall p_j \in P$ і значення цільової функції (8.3). Кількість виконаних ітерацій покласти 1.
3. сформувати нове нечітке розбиття $R'(A) = \{A_k \mid A_k \subseteq A\}$ множини об'єктів кластеризації A на c непусті нечіткі кластери, які характеризуються сукупністю функцій належності $\mu_k^1(a_i), \forall k \in \{2,..., c\}, \forall a_i \in A$, що визначаються за формулою:

$$\mu_k^1(a_i)=\left(\sum_{l=1}^{c}\left(\frac{\left(\sum_{j=1}^{q}(x_i^j-v_k^j)^2\right)^{\frac{1}{2}}}{\left(\sum_{j=1}^{q}(x_i^j-v_l^j)^2\right)^{\frac{1}{2}}}\right)^{\frac{2}{m-1}}\right)^{-1},\ \forall k\in\{2,...,c\},\forall a_i\in A$$

4. При цьому, якщо для деякого $k\in\{2,...,c\}$ і деякого $a_i\in A$ значення $\sum_{j=1}^{q}(x_i^j-v_k^j)^2=0$, тоді для відповідного нечіткого кластеру $\mathrm{A_k}$ беремо $\mu_k^1(a_i)=1$, а для інших кластерів $A_l(\forall l=\{2,...,c\},l\neq k)$ беремо $\mu_l^1(a_i)=0$. Якщо ж таких значень $k\in\{2,...,c\}$ виявиться декілька для $a_i\in A$, тоді евристично беремо $\mu_{\mathrm{k}}^1(\mathrm{a_i})=1$ для меншого з них, а для інших $\mu_l^1(a_i)=0$.
5. Для нового нечіткого розбиття $R'(A)=\{A_k\mid A_k\subseteq A\}$ за формулою (8.2) розраховуємо центри нечітких кластерів і значення цільової функції (8.3).
6. Якщо кількість виконаних ітерацій більше за s або модуль різниці між попереднім і новим значенням цільової функції менше за $\varepsilon\in R_+$, тоді в якості результату прийняти нечітке розбиття $R'(A)=\{A_k\mid A_k\subseteq A\}$ і завершити виконання алгоритму. Інакше, вважати поточним розбиттям $R(A)=R'(A)$ і перейти на крок 3, збільшивши на 1 кількість виконаних ітерацій.

В результаті виконання алгоритм зводиться до деякого локально-оптимального розбиття $R^*(A)$, яке описується сукупністю функцій належності $\mu_k(a_i)$, а також центрами нечітких класів $v_k=(v_k^1,v_k^2,...,v_k^q)$.

8.4. Виконання алгоритму FCM в системі MATLAB

Функція fcm може бути викликана в одному з наступних форматів:

$[center,U,obj\_fcn]=fcm(data,cluster\_n)$

або

$[center,U,obj\_fcn]=fcm(data,cluster\_n,options)$.

Вхідними аргументами цієї функції є

- data: матриця початкових даних D, і-тий рядок якої являє собою інформацію про об’єкт нечіткої кластеризації $a_i\in A$ у формі вектора $x_i=\{x_i^1,x_i^2,...,x_i^q\}$;

- cluster_n: число шуканих кластерів $c\in N,c>1$.

Вихідними аргументами цієї функції є

- center: матриця центрів шуканих нечітких кластерів, кожний рядок якої являє собою координати центру одного з нечітких кластерів в формі вектора $v_k = (v_k^1, v_k^2, ..., v_k^q)$;

- U: матриця значень функцій належності шуканого нечіткого розбиття $\mu_k(a_i), \forall k \in \{2,...,c\}, \forall a_i \in A$;

- obj_fun: значення цільової функції (3) на кожній з ітерацій роботи алгоритму.

Функція fcm() може бути викликана з додатковими аргументами options, які введенні для управління процесом кластеризації, а також для зміни критерію останова роботи алгоритму і/або відображення інформації на екрані монітора.

Ці додаткові аргументи мають наступні значення:

- option (1): експоненційна вага m для розрахунків матриці нечіткого розбиття U (за замовченням $m = 2$);
- option (2): максимальне число ітерацій s (за замовченням це значення дорівнює 100);
- option (3): параметр збіжності алгоритму ε (за замовченням це значення дорівнює 0.00001);
- option (4): інформація про поточну ітерацію, яка відображається на екрані монітора (за замовченням, це значення 1).

Якщо будь-яке зі значень додаткових аргументів дорівнює NaN (не число), тоді для цього аргументу використовується значення за замовченням.

8.5. Приклад реалізації алгоритму

Завдання 1. В якості прикладу застосування нечіткої кластеризації розглянемо множину даних, які містяться в системі MATLAB і використовуються в якості текстової сукупності об’єктів нечіткої кластеризації. Ці дані являють собою матрицю D розмірності 140x2 і містяться у файлі fcmdata.dat, який поставляється разом зі MATLAB В даному випадку матриця D відповідає 140 об’єктам, для кожного з який виконане вимірювання за двома ознаками, що є дуже зручним для візуалізації результатів нечіткої кластеризації в двовимірному просторі на площині.

1. Для візуалізації цих даних слід виконати наступні команди:

Load fcmdata.dat

plot(fcmdata(:,1), fcmdata(:,2), ‘o’)

На екрані з’явиться графічне зображення, яке представлене на рис. 8.1.

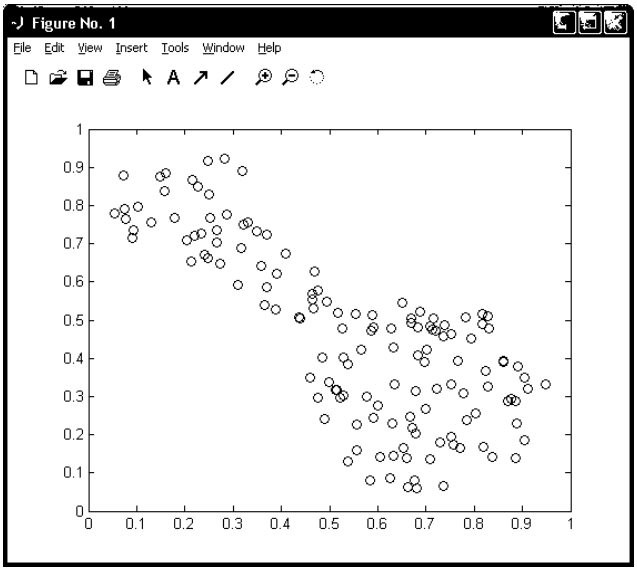

Рис.8.1 Зображення точок матриці D з файлу fcmdata.dat

2. Далі слід викликати функцію fcm, наприклад, з наступним форматом:

[center, U, obj_fcn]=fcm(fcmdata, 2)

Потім слід подивитись результати виконання процедури нечіткої кластеризації:

- координати центрів класів, тобто, матрицю center;
- належність кожної сукупності даних до класів – матрицю U;
- значення функції цілі – obj_fcn.

Взагалі, процедуру нечіткої кластеризації можна записати у вигляді командного М-файла з наступним текстом:

```
load fcmdata.dat
plot(fcmdata(:,1), fcmdata(:,2),’o’)
[center, U, obj_fcn]= fcm(fcmdata, 2);
maxU=max(U);
index1 = find(U(1,:)== maxU);
index2 = find(U(2,:)== maxU);
line(fcmdata(index1,1), fcmdata(index1,2),’linestyle’,’none’,…
‘marker’, ‘x’, ‘color’, ‘g’);
line(fcmdata(index2,1), fcmdata(index2,2),’linestyle’,’none’,…
‘marker’, ‘x’, ‘color’, ‘r’);
hold on
plot( center(1,1), center(1,2),’ko’, ‘markersize’,10, ‘LineWidth’, 2)
plot( center(2,1), center(2,2),’ko’, ‘markersize’,10, ‘LineWidth’, 2)
```

Результатом цієї програми буде розбиття даних на два кластера, візуалізація якого зображена на рис.8.2.

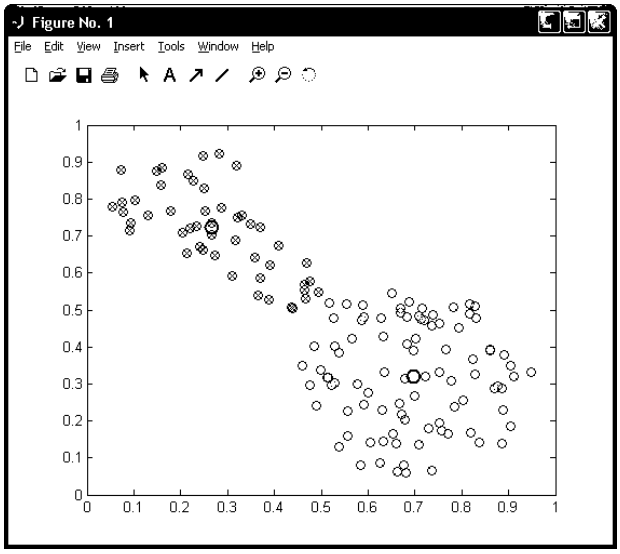

Рис. 8.2. Результат роботи програми нечіткої кластерізації

3. Після роботи програми в системі MATLAB можна перевірити значення матриць center та U, набравши їх назву в командному рядку і натиснувши Enter.

4. Крім того, в програмі можна використати наступний формат запису:

[center, U, obj_fcn]= fcm(fcmdata, 2, [2.5 1000 0.000001 1]);

В цьому випадку експоненційна вага 2.5, максимальне число ітерацій 1000, параметр збіжності $\varepsilon = 0.000001$. Порівняльний аналіз показує практичну ідентичність графіків – результатів використання обох форматів функції fcm, що дозволяє зробити висновок про відповідність отриманих результатів нечіткої кластеризації.

5. Для розв’язування задачі нечіткої кластеризації в системі MATLAB можна використовувати графічний інтерфейс, який викликається за допомогою команди **findcluster**. Ця програма може використовувати або метод с-середніх або метод субтрактивної кластеризації (substractive clustering, який викликається і окремо за допомогою команди **subclust**). Останній використовується тоді, коли не можна заздалегідь встановити число кластерів с на кроці 6.

Формат виклику графічного інтерфейсу: **findcluster або findcluster(‘file.dat’)**.

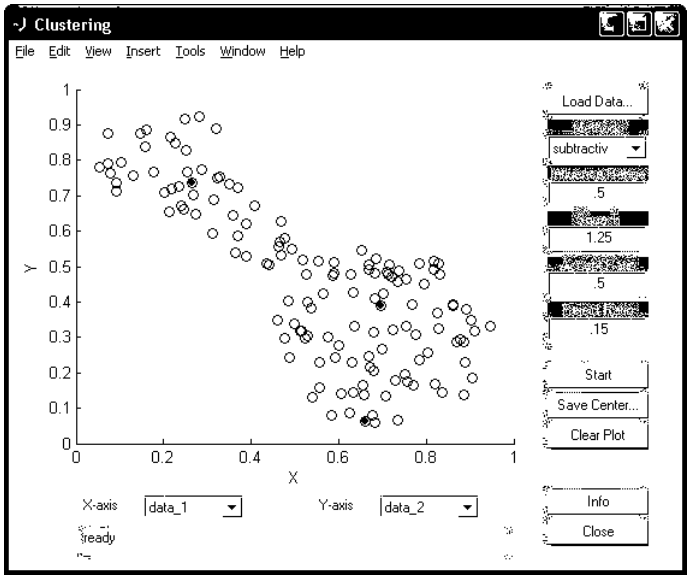

Рис 8.3. Вікно роботи графічного інтерфейсу нечіткої кластеризації для алгоритму субтрактивної кластеризації

В даному вікні можна завантажити файл даних Load Data, обрати метод кластеризації Methods, обрати необхідні значення параметрів і натиснути кнопку Start.

6. Нехай, заздалегідь невідома кількість кластерів c. Тоді слід використати метод субтрактивної кластеризації. Ідея цього методу полягає у тому, що кожна точка даних пропонується в якості центра потенційного кластеру. Далі слід вирахувати деяку міру можливості кожної точки даних представляти центр кластеру. Ця кількісна міра основана на оцінці густини точок навколо відповідного центра кластера.

Цей алгоритм, який є узагальненням методу кластеризації Р. Ягера, заснований на виконанні наступних кроків:

1) вибрати точку даних з максимальним потенціалом для представлення центру першого кластеру
2) забрати всі точки даних в околі центру даного кластеру, величина якої задається параметром **radii**, щоб визначити наступний нечіткий кластер і координати його центру.

Далі, ці дві процедури повторюються до тих пір, допоки всі точки даних не будуть лежати в границях околів радіуса **radii** навколого шуканих кластерів.

Функція командного рядка

[C, S] = subclust (X, radii, xBounds, options)

знаходить центри таких кластерів.

При цьому матриця X містить об’єкти кластеризації, кожний рядок якої відповідає координатам окремої точки даних. Параметр radii являє собою вектор, компоненти якого приймають значення з інтервалу [0,1] і задають діапазон розрахунку центрів кластерів по кожній з розказ вимірювань об’єктів. Робиться припущення, що всі дані знаходяться в деякому гіперкубі. Взагалі, малі значення параметрів radii призводять до знаходження малого числа

великих по кількістю точок кластерів. Найкращих результатів можна очікувати при значенні radii між 0.2 і 0.5.

Аргумент xBounds являє собою матрицю розміру (2xq), яка визначає засіб відображення матриці даних X в деякому одиничному гіперкубі. Тут q – кількість ознак. Цей аргумент є необов'язковим, якщо матриця X вже нормована. Перший рядок цієї матриці містить мінімальні значення інтервалу вимірювання кожної ознаки, а другий рядок – максимальне значення вимірювання кожної ознаки.

Для зміни значень, які встановлені по замовченню, можна використати параметр options, компоненти вектора якого можуть приймати наступні значення:

- options(1) = guashFactor – параметр, який використовується в якості коефіцієнту для множення значень radii з ціллю зменшення впливу потенціалу граничних точок, які розглядаються як частина даного кластеру (за замовченням це значення дорівнює 1.25);
- options(2) = acceptRatio – параметр, який встановлює потенціал як частину потенціалу першого кластеру, вище якого інша точка даних не може розглядатся в якості центра іншого кластеру (за замовченням це значення 0.5);
- option(3) = rejectRation - параметр, який встановлює потенціал як частину потенціалу першого кластеру, нижче якого інша точка даних не може розглядатся в якості центра іншого кластеру (за замовченням це значення 0.15);
- options(4) = verbose – якщо значення цього параметра не дорівнює 0, тоді на екран монітору виводиться інформація про виконання процесу кластеризації (за замовченням це значення 0).

Функція subclust повертає матрицю C значень координат центрів нечітких кластерів. При цьому кожний рядок цієї матриці містить координати одного центру кластеру. Вектор S містить σ-значень, які визначають діапазон впливу центра кластеру по кожній з розглянутих ознак. При цьому, всі центра кластерів мають однакову множину σ-значень.

Для **прикладу** розглянемо наступну послідовність команд:

Load fcmdata.dat

[C, S] = subclust(fcmdata, [0.5 0.5], [], [1.25 0.5 0.15 1])

Для кожної ознаки вводяться радіуси околів – 0.5 і 0.5.

Як видно з рис. 3. для даної сукупності даних функція suclust має знайти три кластери.

Для розв'язування задачі субтрактивної кластеризації може використовуватися і розглянутий раніше графічний інтерфейс, який викликається командою **findcluster**.

Таким чином, система MATLAB дозволяє розв'язувати задачі нечіткої кластеризації двома способами: за допомогою функцій командного рядка і за допомогою графічного інтерфейсу кластеризації. Результати нечіткої кластеризації мають наближений характер і мають служити для попередньої структуризація вхідної (початкової) інформації про систему, що вивчається. Тобто, провівши кластерізацію, можна отримати і зберегти знання про систему у вигляді структуризації початкової інформації.

8.6. Завдання для самостійної роботи

1. Виконати нечітку кластеризації набору точок в двовимірному просторі ознак: (1,2); (2,1); (1, 2); (3,2); (2,3); (1,6); (2,5); (3,6); (5,5); (5,6); (7,5); (7,6). Розділити їх на 3 кластери. Відповідь можна подивитися на рис 8.4.

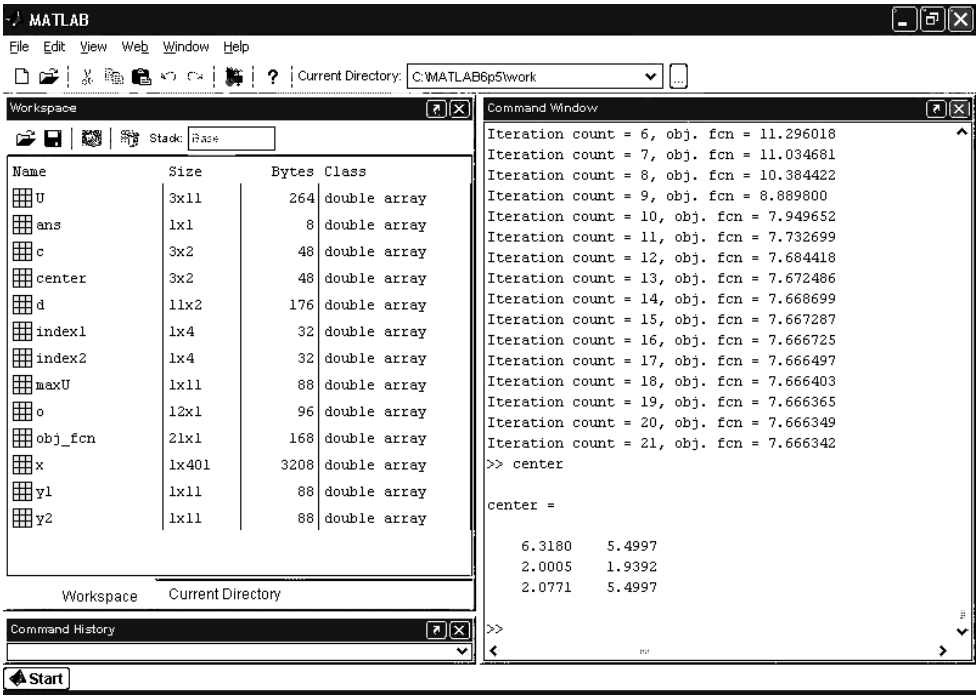

Рис 8.4. Точки і центри кластерів завдання 1

2. Виконати нечітку кластерізацію 10 фірм, які можуть в майбутньому опинитися в різних зонах своєї діяльності на ринку (2, 3 або 4 зони). Для цього слід використовувати дві ознаки стану фірм – середній дохід і середньо-квадратичне відхилення можливого доходу, тобто, ризикованість. Відомі наступні дані про фірми:

Фірма	Показники					
1	Можливий дохід	100	120	300	250	200
	Імовірність	0,1	0,3	0,2	0,2	0,2
2	Можливий дохід	120	400	200	150	300
	Імовірність	0,2	0,4	0,1	0,2	0,1
3	Можливий дохід	200	250	260	150	300
	Імовірність	0,3	0,3	0,2	0,1	0,1
4	Можливий дохід	120	100	200	150	130
	Імовірність	0,2	0,4	0,2	0,2	0,2
5	Можливий дохід	520	400	250	450	300
	Імовірність	0,6	0,1	0,1	0,1	0,1
6	Можливий дохід	520	400	200	450	300
	Імовірність	0,2	0,4	0,1	0,2	0,1

7	Можливий дохід	220	300	200	150	300
	Імовірність	0,2	0,3	0,2	0,2	0,1
8	Можливий дохід	100	300	200	150	300
	Імовірність	0,2	0,2	0,3	0,1	0,2
9	Можливий дохід	230	400	200	250	200
	Імовірність	0,1	0,3	0,2	0,2	0,2
10	Можливий дохід	220	120	200	350	400
	Імовірність	0,1	0,5	0,1	0,2	0,1

Яка можлива економічна інтерпретація цих зон?