Разпознаване на образи

1. Основни понятия

Способността „да се разпознава“ се счита за основно свойство на човешките същества, а може би и на останалите живи организми.

Под понятието ***образ*** се разбира описанието на съответния обект. Във всеки един момент, когато не спим, ние изпълняваме разпознаване на образи. Можем да различим свой познат в тълпа от много хора и да разберем какво говори; да разпознаем гласа на познат; да прочетем ръкописен текст; да различим усмивка от злобна гримаса.

Могат да се отделят два основни типа разпознаване на образи на базата изпълнение на самото разпознаване:

1.1. разпознаване на конкретни обекти – можем да разпознаваме символи, рисунки, знаци, музика и обекти от нашето обкръжение. Използва се слух или зрение и процесът още се нарича ***сензорно разпознаване***.

1.2. разпознаване на абстрактни обекти – процесът е наречен още ***понятийно разпознаване*** – със затворени очи и запушени уши да се определят стари доводи (разпознаят) или да се реши задача.

Съществуват две основни направления в процеса на разпознаване на образи:

1. Изследване и изучаване на способностите за разпознаване, които притежават човешките същества и други живи организми.

2. Развитие на теорията и методите за създаване на устройства, предназначени за решаване на отделни задачи за разпознаване на образи в различни приложни области.

Ние ще се интересуваме от второто направление, което касае техническата страна при построяването на автоматични системи за разпознаване на образи.

По-просто казано, разпознаването на образи може да се определи като отнасяне на изходните данни към един или друг клас с помощта на отделяне на съществените признаци или свойства, характеризиращи тези данни, от общата маса несъществени детайли.

Системата за разпознаване на символи представлява система за разпознаване на образи – като изходни данни се подават оптични сигнали, а системата идентифицира имената на символите.

Образ – това е описание на произволен елемент като представител на съответен клас образи.

Когато множество образи се разделя на непресичащи се класове е удобно да се използва автоматично устройство за сортиране на отделните елементи по класове – сортиране на монети, идентифициране на регистрационни табели, класифициране на звуци.

2. Основни задачи, възникващи при разработване на системи за разпознаване на образи

Образ – описание на конкретен обект чрез набори от признаци, характеризиращи отделни негови свойства.

Клас от образи – множество от образи, притежаващи общи признаци.

Разпознаване на образ означава отнасянето му към определен клас.

2.1. Първа задача – представяне на изходните данни

Образите се описват най-лесно с помощта на вектори:

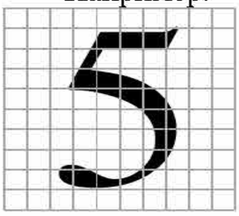
$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$ или с транспонирания вектор $x = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$,

където n – брой на признаците (характеристиките).

Практически не се използват всички признаци на обекта, а се минимизират на база на тяхната информативност.

x_i – стойност на i-тия признак. Тази стойност може да бъде двоична $x_i = \{0, 1\}$ или количествена $x_i = [a, b]$.

Например:



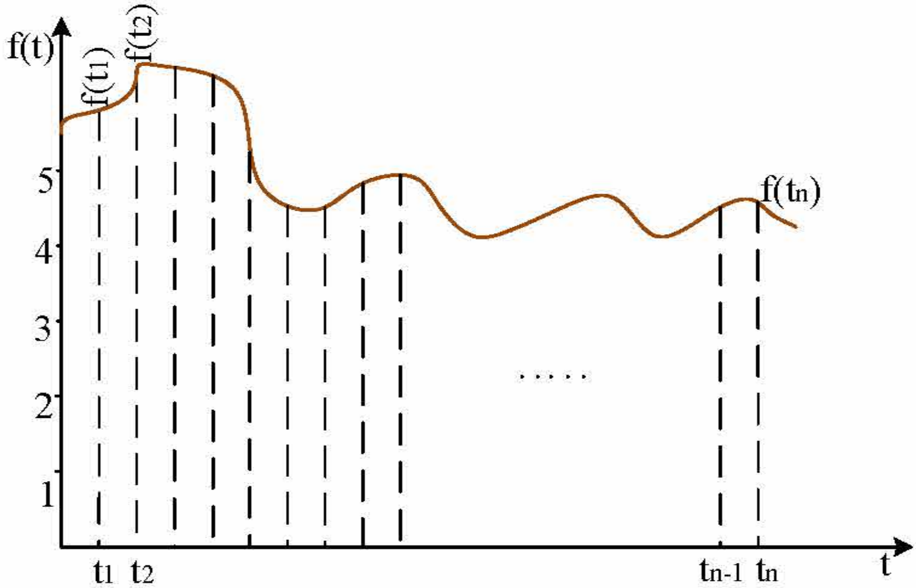
$x_i = 1$, ако в i-тото квадратче минава линия от обекта (цифрата).

$x_i = 0$, ако в i-тото квадратче **не** минава линия от обекта.

$x = (x_1, x_2, \ldots, x_n)$

$x_i = 0$ или $x_i = 1$, $i = 1 \ldots n$

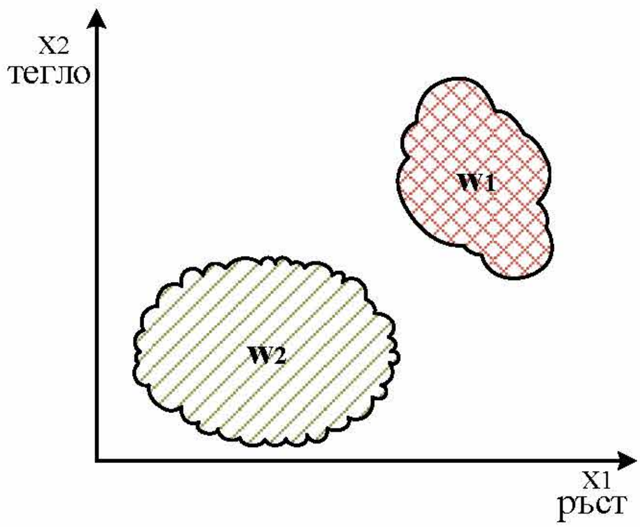
Имаме графика на непрекъсната функция (звуков сигнал) $f(t)$ във времето t. Ако стойностите на функцията се определят в дискретни моменти $t_1 \div t_n$, $x = (x_1, x_2, \ldots, x_n)$, $x_i = f(t_i)$



Векторите на образите съдържат подлежаща на измерване информация за образите. Процесът на измерване, на който се подлагат обектите от определен клас образи, може да се разглежда като процес на кодиране, заключаващ се в присвояване на стойност на всяка отделна характеристика $\{x_i\}$. Често измерванията дават като резултат числа и е по-нагледно и полезно да се разглеждат векторите като точки в n-мерното пространство.

Множеството образи, принадлежащи на един и същ клас, се изобразява в съвкупност от точки, разсеяни в дадена област от това n-мерно пространство.

Имаме два класа w_1 и w_2, които съответстват на две групи спортисти. Всеки образ се характеризира с резултатите от две измервания: ръст и тегло. Следователно, векторите имат вида $x = (x_1, x_2)$, където x_1 е ръст, а x_2 е тегло. Всеки вектор може да се разглежда като точка от двумерното пространство.



Както се вижда двата класа w_1 и w_2 са непресичащи се множества, което на практика се постига много трудно. На практика е трудно да се подберат параметрите така, че да се получат строго непресичащи се множества за отделните класове.

2.2. Втора задача – отделяне на характерните признаци или свойства от получените изходни данни и намаляване на размерността на векторите.

Общите признаци, характерни за всички образи от даден клас, представляват характерните свойства на класа или признаците на класа.

Признаците, характеризиращи различията между отделните класове могат да се интерпретират като междукласови признаци.

Общите за всички класове признаци не носят полезна информация от гледна точка на разпознаването на образи и могат да не се вземат под внимание.

Изборът на признаци е една от най-важните задачи, свързани със създаването на системи за разпознаване.

Най-лесният вариант е, ако в резултат от различни измервания се получава пълния комплект от признаци за всички класове. Тогава самото разпознаване и класификация на образите няма да предизвикат затруднения. Автоматичното разпознаване се свежда до процес на просто съпоставяне или процедура на просто преглеждане на таблица.

На практика такъв идеален вариант не съществува. От изходните данни обикновено могат да бъдат извлечени някои от признаците, а векторите могат да се минимизират с помощта на преобразувания, гарантиращи минимална загуба на информация.

2.3. Трета задача – определяне на решаващите функции за идентификация и класификация.

След като данните, събрани за подлежащите на разпознаване образи, са представени с точки в съответното пространство или с вектори, машината или системата трябва да определи на какъв клас образи съответстват тези данни.

Нека системата е предназначена за M класа, означени с $w_1, w_2, \ldots, w_M$. Пространството на образите може да се смята, че се състои от M области, всяка от които съдържа само точки, съответстващи на образи от един клас. По този начин задачата за разпознаване може да се разглежда като задача за определяне на границите на тези M области на базата на регистрираните/измерените вектори.

В случаите, когато за разпознаваните образи има малка априорна информация, решаващите функции могат да бъдат определени на базата на тази информация. Ако за образите има само качествени данни, то могат да бъдат определени допустими приближения за вида на решаващата функция, но границите, определени от нея, да се различават чувствително и да се налагат множество корекции. Поради това, при липса на априорна информация, е добре в автоматичната система да се използват обучаващи процедури.

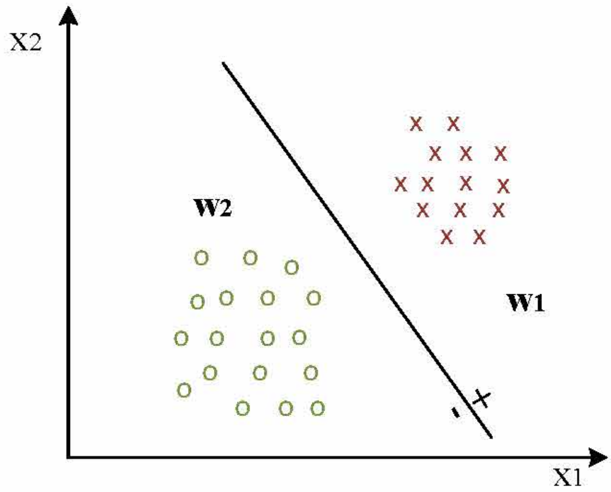
- Да се изберат решаващите функции произволно;
- Чрез итеративни стъпки на обучение тези функции да се модифицират до оптимален или приемлив вариант.

Методи за разпознаване на образи
Решаващи функции

1. Въведение

Основна задача на системите за разпознаване на образи е да се определи критерий, според който отделните образи да бъдат коректно класифицирани в съответния им клас. Един от методите, намерили приложение в решаването на тази задача, е методът, базиран на решаващите функции.

Нека имаме следната ситуация:



Имаме образи, принадлежащи на два класа w_1 и w_2.

Както се вижда от графичното представяне – двете групи от образи могат да бъдат разделени с права линия. Уравнението на тази разделяща права има вида:

$$d(x) = w_1 x_1 + w_2 x_2 + w_3 = 0$$

w_i – параметри, x_1 и x_2 – променливи.

От графичното представяне се вижда, че при заместване в $d(x)$ с произволен образ x от класа w_1 ще се получи положителен резултат. Отрицателни стойности ще получаване за $d(x)$ за всеки образ от класа w_2.

По такъв начин $d(x)$ може да се използва като **решаваща функция** или **разделяща/дискриминантна функция**. Ако разглеждаме образ x с неизвестна все още класификация, то може да се твърди, че x принадлежи на w_1, ако $d(x) > 0$ или на класа

w_2, ако $d(x)<0$. Ако образът x лежи на граничната линия, имаме случай на неопределеност и $d(x)=0$.

Този метод е приложим и работи коректно и за по-голям брой класове от 2. Успешното му приложение зависи от два фактора:

1. видът на функцията $d(x)$;
2. практическата възможност да бъдат определени коефициентите в $d(x)$.

Първият фактор е непосредствено свързан с геометричните свойства на разглежданите класове. Възможна и очевидна е ситуация, при която за разделяне на групите/класовете от образи са необходими граници, далеч по-сложни за описание от права линия. Ако размерността е по-голяма от 3, то и зрителната ни представа спира да помага за определяне на тези области. В такива случаи единствен разумен подход е използването на аналитични процедури. За съжаление, когато няма априорна информация е много сложно и може само емпирично да се направи оценка за ефективността на избраната решаваща функция.

След като бъде определена функцията трябва да бъдат определени коефициентите. За решаването на този проблем се използват различни методи чрез адаптивни схеми или обучаващи процедури.

2. Линейни решаващи функции

Показаният в точка 1 най-прост вариант на двумерна решаваща функция може лесно да се обобщи за n-мерния случай, като се използва формулата:

$$d(x)=w_1x_1+w_2x_2+\ldots+w_nx_n+w_{n+1}=w_0'x+w_{n+1}.$$

x – вектор с характеристики на образите, $w_0=(w_1, w_2, \ldots, w_n)'$ – тегловен или параметричен вектор.

Общоприето е във всички вектори на образи да се добавя компонент 1 след последния показател или признак и за линейната решаваща функция имаме:

$$d(x)=w'x,$$

където $x=(x_1, x_2, \ldots, x_n, 1)'$ и $w=(w_1, w_2, \ldots, w_n, w_{n+1})'$ са допълнените вектори на образите и параметрите съответно.

Предполага се, че в случай на разделяне в два класа, решаващата функция $d(x)$ притежава следните свойства:

$$d(x)=w'x=\begin{cases}>0, & \text{ако } x\in w_1\\ <0, & \text{ако } x\in w_2\end{cases}$$

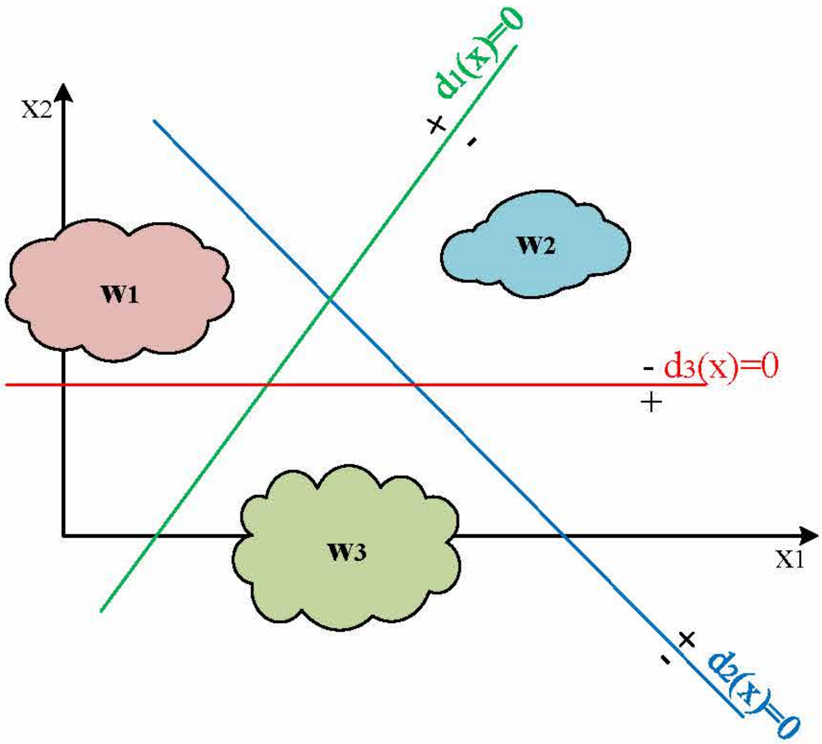
Да разгледаме по-общия случай разделяне на $w_1, w_2, \ldots, w_M$ класа (повече от два класа).

2.1. Случай 1 – всеки клас се разделя от всички останали класове с една разделяща повърхност. В този случай съществуват M решаващи функции със следните свойства:

$$d_i(x)=w_i'x=\begin{cases}>0, & \text{ако } x\in w_i\\ <0, & \text{ако } x\notin w_i\end{cases} \text{ за } i=1, 2, \ldots, M,$$

където $w_i=(w_{i1}, w_{i2}, \ldots, w_{in}, w_{i\,n+1})'$ е тегловният вектор на i-тата решаваща функция.

Пример:



w_1, w_2, w_3 – три класа и всеки от тях можем да отделим от другите два с една разделяща права.

Ако даден образ x принадлежи на w_1, то за него следва, че: $\begin{matrix} d_1 > 0 \\ d_2 < 0 \\ d_3 < 0 \end{matrix}$

Ако даден образ x принадлежи на w_2, то за него следва, че: $\begin{matrix} d_1 < 0 \\ d_2 > 0 \\ d_3 < 0 \end{matrix}$

Ако даден образ x принадлежи на w_3, то за него следва, че: $\begin{matrix} d_1 < 0 \\ d_2 < 0 \\ d_3 > 0 \end{matrix}$

Ако $d_i(x) > 0$ за повече от една стойност на *i*, тази схема на класификация няма решение. Това важи за $d_i(x) < 0$ за всяко *i*. И както се вижда, имаме четири области на неопределеност (ОН).

Ако имаме конкретни числови коефициенти, например:

$$\left| \begin{array}{l} d_1(x) = -x_1 + x_2 \\ d_2(x) = x_1 + x_2 - 5 \\ d_3(x) = -x_2 + 1 \end{array} \right.$$

Следователно, трите разделящи линии имат следните уравнения:

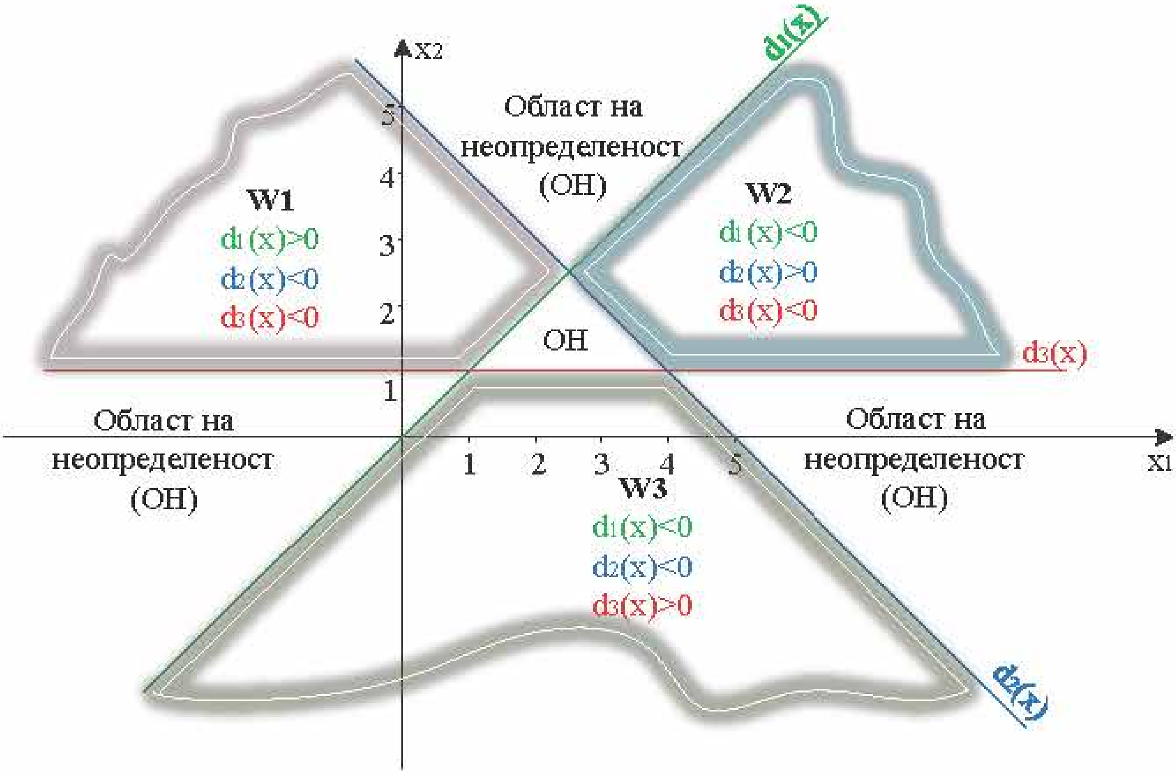
$$\begin{cases} -x_1 + x_2 = 0 \\ x_1 + x_2 - 5 = 0 \\ -x_2 + 1 = 0 \end{cases}$$

за права d_1: при $x_1 = 0 \implies x_2 = 0$ и при $x_1 = 1 \implies x_2 = 1$. Правата минава през точки с координати (0,0) и (1,1) . Две точки са достатъчни за построяване на права!

за права d_2: при $x_1 = 0 \implies x_2 = 5$ и при $x_2 = 0 \implies x_1 = 5$. Правата минава през точки с координати (0,5) и (5,0)

за права d_3: при $x_1 = 0 \implies x_2 = 1$ и при $x_1 = 1 \implies x_2 = 1$. Правата минава през точки с координати (0,1) и (1,1).

Ето как изглеждат разделящите линии в този пример:

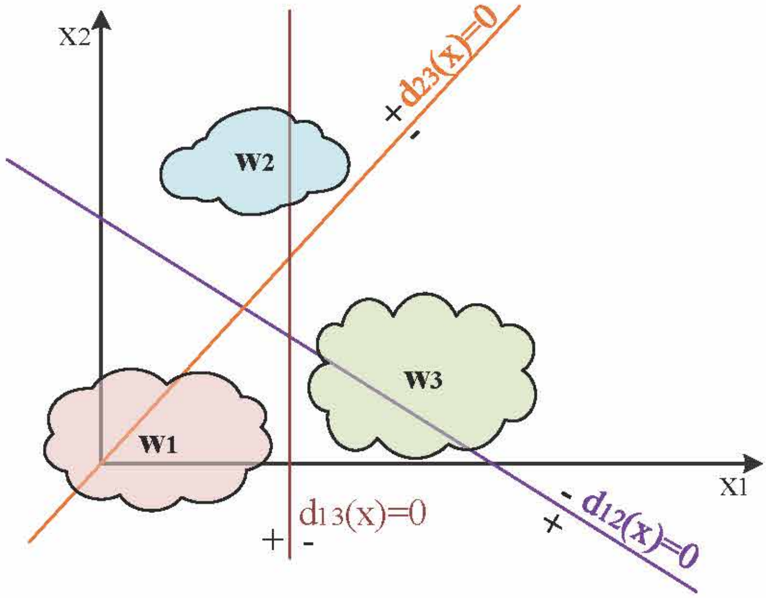


Ако трябва да се класифицира образ $x = (6,5)$, заместваме в трите решаващи функции:

$$\begin{cases} d_1(x) = -x_1 + x_2 = -6 + 5 = -1 < 0 \\ d_2(x) = x_1 + x_2 - 5 = 6 + 5 - 5 = 6 > 0 \quad \Rightarrow \quad \text{образът е от клас } w_2. \\ d_3(x) = -x_2 + 1 = -5 + 1 = -4 < 0 \end{cases}$$

2.2. Случай 2 – всеки два класа се разделят помежду си с разделяща повърхнина/равнина. В този случай съществуват $M(M-1)/2$ разделящи повърхности. Решаващите функции имат вида $d_{ij}(x) = w_{ij}x$ и притежават следните свойства: ако образът x принадлежи на класа w_i, то $d_{ij}(x)>0$ за всички $j \neq i$ и освен това $d_{ji}(x) = -d_{ij}(x)$.

Пример:



Имаме w_1, w_2, w_3 – три класа. Всяка разделяща е за точно два класа:

$d_{13}(x)$ – за w_1 и w_3;

$d_{12}(x)$ – за w_1 и w_2;

$d_{23}(x)$ – за w_2 и w_3.

Нека имат следните числови коефициенти:

$d_{12}(x) = -x_1 - x_2 + 5$

$d_{13}(x) = -x_1 + 3$

$d_{23}(x) = -x_1 + x_2$

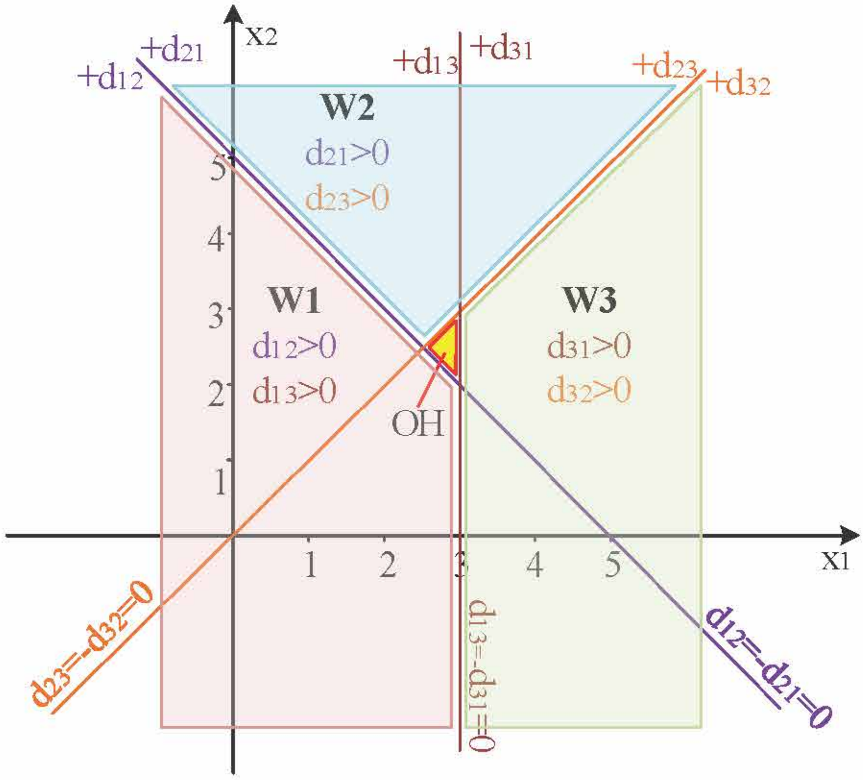
Граничните линии получаваме като приравним на 0.

$d_{12}(x) = -x_1 - x_2 + 5 = 0$

$d_{13}(x) = -x_1 + 3 = 0$

$d_{23}(x) = -x_1 + x_2 = 0$

Можем да начертаем конкретните прави в координатната система и имаме следната ситуация:



Нека имаме образ $x=(4,3)$.

$d_{12}(x)=-x_1-x_2+5=-4-3+5=-2<0$ $\qquad\qquad d_{21}(x)=2>0$

$d_{13}(x)=-x_1+3=-4+3=-1<0$ $\qquad \Rightarrow \qquad d_{31}(x)=1>0$

$d_{23}(x)=-x_1+x_2=-4+3=-1<0$ $\qquad\qquad d_{32}(x)=1>0$

$\Rightarrow\ d_{3i}(x)>0$ и не попада в областта на неопределеност $\Rightarrow$ принадлежи на клас w_3.

2.3. Случай 3 – съществуват M решаващи функции $d_k(x)=w_k'x,\ k=1,2,\ldots,M$, такива, че ако образът $x\in w_i$, то $d_i(x)>d_j(x)$ за всяко $j\neq i$. Този случай е разновидност на случай 2, тъй като можем да приложим:

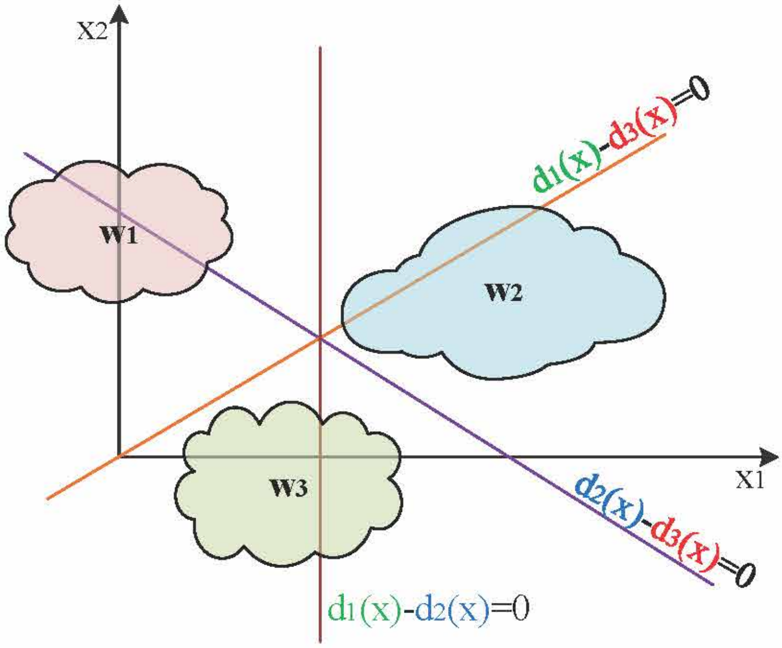
$d_{ij}(x)=d_i(x)-d_j(x)=(w_i-w_j)'x=w_{ij}'x$.

Ако $d_i(x)>d_j(x)$ за всяко $j\neq i$, то $d_{ij}(x)>0$.

Пример:

Границата между w_i и w_j се определя от тези значения на x, при които е изпълнено равенството: $d_i(x)=d_j(x)$ или $d_i(x)-d_j(x)=0$. Следователно, при извеждането на уравнението на разделящата за w_i и w_j се използват заедно стойностите на решаващите функции $d_i(x)$ и $d_j(x)$.

Прост вариант при $M=3$:



За всички образи $x \in w_1$ трябва:

$$\begin{cases} d_1(x) > d_2(x) \\ d_1(x) > d_3(x) \end{cases}$$

или да са в положителните зони на $d_1(x) - d_2(x) > 0$ и $d_1(x) - d_3(x) > 0$.

Нека имаме следните конкретни решаващи функции:

$$\begin{cases} d_1(x) = -x_1 + x_2 \\ d_2(x) = x_1 + x_2 - 1 \\ d_3(x) = -x_2 \end{cases}$$

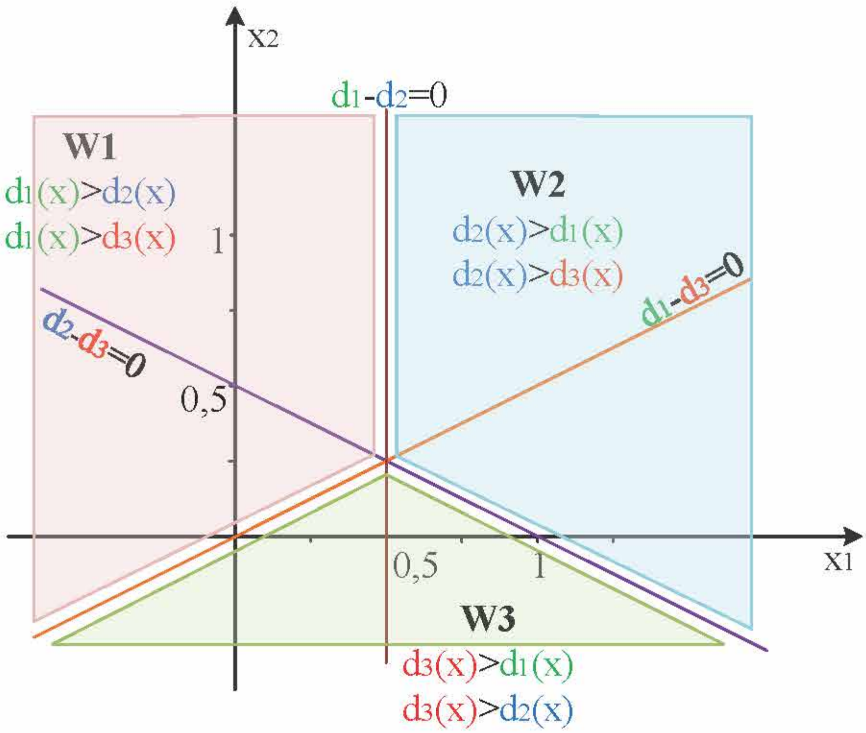
Разделящите граници са съответно:

$$d_1(x) - d_2(x) = -x_1 + x_2 - x_1 - x_2 + 1 = -2x_1 + 1 = 0$$

$$d_1(x) - d_3(x) = -x_1 + x_2 + x_2 = -x_1 + 2x_2 = 0$$

$$d_2(x) - d_3(x) = x_1 + x_2 - 1 + x_2 = x_1 + 2x_2 - 1 = 0$$

Графичното представяне на разделящите функции:



Няма област на неопределеност!

Ако имаме образ $x = (1,1)$:

$d_1(x) = -x_1 + x_2 = -1 + 1 = 0$

$d_2(x) = x_1 + x_2 - 1 = 1 + 1 - 1 = 1 > 0$

$d_3(x) = -x_2 = -1 < 0$

$\Rightarrow d_2(x) > d_1(x)$ и $d_2(x) > d_3(x)$ $\Rightarrow x \in w_2$.