Математички Билтен
18 (XLIV)
1994 (103-114)
Скопје, Македонија

ISSN 0351-336X

ПРОБЛЕМОТ НА ОБУЧУВАЊЕ НА НЕВРОНСКА МРЕЖА КАКО СПЕЦИЈАЛНА ЗАДАЧА НА ЦЕЛОБРОЈНОТО ЛИНЕАРНО ПРОГРАМИРАЊЕ

Ѓорѓи Јованчевски, Стево Божиновски

При разгледувањето на проблемот на обучување на невронска мрежа во просторот на тежини R^m утврдената целобројност на обучувањето овозможува премин во наставниот простор од целобројни позитивни вектори p на евклидскиот n-димензионален простор R^n и со избор на соодветен критериум на оптималност, проблемот се сведува на специјална задача на минимизација од областа на целобројното линеарно програмирање.

1. Вовед

Обучувањето на невронска мрежа се состои во одредување (наоѓање) на асоцијативна меморија (мемориска матрица на тежини) W*, која што ги памти (препознава) сите вектори на даденото обучувачко множество S.

Обучувањето на мрежата се изведува со тренажен процес на последователни обиди, во кој на мрежата ѝ се покажува секој лик од обучувачкото множество S и се испитува дали таа го препознава ликот или не. Тренажниот процес започнува од состојба "tabula rasa", кога мрежата не поседува никакво знаење за ликовите, т.е. нејзината меморија се смета за празна.

Формулација на проблемот

Нека е дадено обучувачкото множество S од n ликови претставени со n во парови линеарно независни ненегативни целобројни вектори $s_j=[s_{ij}]_{m\times 1}$ од R^m. Значи, $S=\{s_j\in R^m \mid j=1,2,\dots,n\}$. Се бара да се одреди множеството W* од мемориски вектори $w_j^*=[w_{ij}^*]_{m\times 1}$, $j\in N=\{1,2,\dots,n\}$, со кое мрежата ќе може да ги препознава сите ликови s_k, $k\in N$. Значи, се бара $W^*=\{w_j^*\in R^m \mid j\in N\}$ така што за секој лик s_k, $k\in N$ се исполнети условите (неравенствата):

$$g_k(w_k^*) > g_k(w_j^*) \text{ за секој } j\neq k, \tag{1.1}$$

каде што

$$g_k(w_j^*) = s_k^T w_j^* - \delta_j, \quad j \in N \tag{1.2}$$

се активностите на соодветните неврони, предизвикани од ликот s_k, а δ_j се прагови на невроните од мрежата.

Според тоа, овде ќе се задржиме на проблемот на асоцијација на парови [4].

Во почетокот имаме $w_j=0$, $j \in N$ и природно е процесот на препознавање да отпочне со ликот s_k за $k=1$.

Ако е исполнет условот

$$g_k(w_k) > g_k(w_j) \text{ за секој } j \neq k, \tag{1.1'}$$

тоа значи дека мрежата го препознава ликот s_k и процесот треба да продолжи за препознавање на наредниот лик s_{k+1}. Обучувањето на мрежата ќе заврши кога мрежата ќе го препознае и последниот лик s_n.

Ако условот (1.1') не е исполнет за барем еден индекс j, тогаш е потребно обучување на мрежата, т.е. извршување на обучувачки обид за памтење на ликот s_k. Ваквиот обид се реализира со некое од познатите правила на учење.

Ние се определуваме за наједноставното, недекрементирачко правило на учење [3],

$$\Delta w_k = c s_k \tag{1.3}$$

каде што c е избрана позитивна константа, наречена брзина на учење.

Со ова правило при ℓ-тиот обид се извршува следнава корекција на меморијата (од ℓ-1-от обид) на невроните,

$$w_j^{(\ell)} = \begin{cases} w_j^{(\ell-1)} + c s_k, & \text{за } j=k \\ w_j^{(\ell-1)} & \text{, за } j \neq k \end{cases}, \quad j \in N$$

и за новите вредности $w_j=w_j^{(\ell)}$, $j \in N$ се проверува дали е исполнет условот (1.1') за s_{k+1} во улога на s_k.

2. Проблемот на целобројност и премин во наставниот простор

При испитување на зависноста на меморијата W на невронската мрежа од брзината на учење c, за различни вредности на c се добиваат различни мемориски вектори. На пример, при $S = \{s_1, s_2, s_3\}$, каде што $s_1 = [1\ 0\ 2]^T$, $s_2 = [2\ 1\ 3]^T$, $s_3 = [3\ 2\ 4]^T$, за четири различни вредности на c добиени се соодветни мемориски вектори w_i, i=1,2,3 содржани во следнава табела:

c	w_1^T	w_2^T	w_3^T
0.17	[3.57;0.00; 7.14]	[4.42; 2.21; 6.63]	[4.59; 3.06; 6.12]
0.50	[10.50;0.00;21.00]	[13.00; 6.50;19.50]	[13.50; 9.00;18.00]
0.83	[17.43;0.00;34.86]	[21.58;10.79;32.37]	[22.41;14.94;29.88]
1.00	[21.00;0.00;42.00]	[26.00;13.00;39.00]	[27.00;18.00;36.00]

Во тренажниот процес, за секој од четирите случаи, беше потребно да се извршат 66 обиди, од кои 43 беа обучувачки. Во секој случај, извршени се ист број обучувачки обиди за секој лик и тоа: за првиот лик 21 обучување, за вториот лик 13 обучувања и за третиот лик 9 обучувања. Ова согледување го наметнува заклучокот дека меморијата на мрежата не може еднозначно да се определи, туку еднозначно е определен најмалиот број на учења на мрежата за препознавање на секој лик од обучувачкото множество. Значи, за да се обучи мрежата, мора за секој лик да се изврши одреден (минимален) број учења (обучување).

Оваа појава се среќава и кај човекот. За да може човекот да запамети некој лик (на објект, на појава, на поим итн.), тој лик мора да го видел, пред тоа, барем еднаш. Колку пати треба да го види човекот некој лик, за да го запамети, зависи од тоа дали тој лик е сличен со некој лик што веќе го познава. Позната е појавата дека два близнаци потешко се препознаваат, односно треба да се видат повеќе пати додека се запамети кој е кој, отколку ако не се работи за близнаци.

Се поставува прашањето дали може и невронската мрежа да се обучи да препознава ликови, ако е познат само бројот на обучувања на ликовите од обучувачкото множество. И уште, дали може да се одреди минималниот број на обучувања на мрежата за секој лик од обучувачкото множество, кој е доволен за да може мрежата да го препознава секој лик.

Во општ случај, за обучувачкото множество $S = \{s_j \in R^n \mid j \in N\}$, нека се $p_1^*, p_2^*, \ldots, p_n^*$ броевите на учења на мрежата за соодветните ликови $s_1, s_2, \ldots, s_n$ по завршувањето на обучувањето. Тогаш меморисиките вектори можат да се изразат преку p_j^* и s_j, $j \in N$,

$$w_j^* = \underbrace{0+cs_j+cs_j+cs_j+\ldots+cs_j}_{p_j^* \text{ пати}} = cp_j^* s_j \quad \text{за } j \in N \tag{2.1}$$

и при тоа $p_j^* \geq 1$, $j \in N$.

Броевите на учења на мрежата за векторите од обучувачкото множество, можеме да ги претставиме со векторот на учење p^*, од n-димензионалниот евклидски простор R^n.

$$p^* = [p_1^* \quad p_2^* \quad \ldots \quad p_n^*]^T \in R^n,$$

каде што p_j^*, $j \in N$ се целобројни и позитивни.

Просторот R^n, кој ги содржи целите позитивни вектори p, се нарекува <u>наставен простор</u> или <u>простор на учителот</u>, [2].

Согласно (2.1), условот (1.1') за препознавањето на векторот s_k, може да се изрази преку векторот p,

$$g_{kk}(p) > g_{kj}(p) \quad \text{за } \forall j \neq k \tag{2.2}$$

каде активноста на невроните (1.2), во зависност од векторот на учење p, кој го одредува актуелното обучување е

$$g_{kj}(p) = cs_k^T s_j e_j^T p - \delta_j, \quad \text{за } j \in N; \tag{2.3}$$

e_j, $j \in N$ се единечните вектори од R^n.

Сега, проблемот на обучување на мрежата може да се искаже на следниот начин: Ако, за актуелниот вектор на учење p, кога на мрежата ќе ѝ се покаже ликот s_k, е исполнето неравенството (2.2), тогаш ликот е препознат, а во спротивно не. Значи, за препознавање на ликовите, потребно е да се одреди целоброен вектор на учење $p \in R^n$.

При препознавање на влезниот лик s_k, велиме дека невронската мрежа ќе го препознае, ако активноста на k-от неврон предизвикана од s_k е поголема од активноста на кој било друг неврон предизвикана од истиот лик. Аналогно, при препознавањето на ликот s_j, активноста на j-от неврон е поголема од активноста на k-от неврон, ако и двете се предизвикани од ликот s_j. Оттука може да се каже дека за препознавање на ликот s_k во однос на ликот s_j, активноста на k-от неврон предизвикана од s_k, треба да биде поголема од активноста на истиот неврон предизвикана од s_j.

Активноста на k-от неврон предизвикана од ликот s_k, согласно (2.3), може да се изрази на следниот начин:

$$g_{kj}(\mathbf{p}) = ch_{kj}e_j^T p - \delta_j, \text{ за } j \in N \tag{2.4}$$

каде $h_{kj} = s_k^T s_j \geq 0$.

Условот (2.2), со оглед на (2.4) и претпоставката сите неврони да имаат ист внатрешен потенцијал, т.е. $\delta = \delta_j$, $j \in N$, го добива следниот облик,

$$h_{kk}p_k - h_{kj}p_j > 0. \tag{2.5}$$

Ако ја воведеме ознаката

$$h_{kj}^T = [0 \ldots 0 \ h_{kk} \ 0 \ldots 0 \ -h_{kj} \ 0 \ldots 0] \tag{2.6}$$

тогаш условот (2.5) може да се изрази како скаларен производ на векторите h_{kj} и $\mathbf{p}$,

$$h_{kj}^T\mathbf{p} > 0. \tag{2.7}$$

Концептот на просторот на обучување се покажува интересен и од математичка гледна точка и во натамошниот текст ќе изложиме некои наши резултати во овој домен.

3. Егзистенција на вектори на учење

Во евклидскиот простор $\mathbb{R}^n$ за произволно даден вектор $a \neq 0$, множеството од вектори

$$Q = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} > 0\}$$

претставува отворен (десен) полупростор определен со хиперрамнината низ координатниот почеток

$$H = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} = 0\}$$

така што $Q \cup H = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} \geq 0\}$ е затворен (десен) полупростор, определен со H, за кој секоја точка од H е гранична.

<u>Лема 3.1</u>. За произволно дадени ненулти вектори $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, нека $Q_1 = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} > 0\}$ и $Q_2 = \{\mathbf{x} \mid \mathbf{b}^T\mathbf{x} > 0\}$. Тогаш $Q_1 \cap Q_2 \neq \emptyset$ ако и само ако $\mathbf{a}^T\mathbf{b} \neq -1$.

<u>Доказ</u>. Без губење од општоста, може да се претпостави дека векторите a и b имаат должина 1 т.е. $\|a\| = \|b\| = 1$. Нека прво $a^Tb \neq -1$. Тогаш, согласно Шварцовото неравенство, е точно дека

$$a^Tb > -1$$

и, на пример за $x=a+b$, имаме

$$a^Tx = a^Ta + a^Tb > 1 + (-1) = 0$$

$$b^Tx = b^Ta + b^Tb > -1 + 1 = 0$$

што значи дека $x \in Q_1 \cap Q_2$. И обратно, нека $Q_1 \cap Q_2 \neq \emptyset$. Да претпоставиме дека $a^Tb=-1$ што повлекува $b=-a$. За $\bar{x} \in Q_1 \cap Q_2$ тогаш мора да биде $a^T\bar{x} > 0$ и $b^T\bar{x} = (-a^T)\bar{x} = (-1)(a^T\bar{x}) > 0$ што претставува противречност.

Од погорната Лема непосредно следат следниве заклучоци:

<u>Заклучок 1</u>. Ако $\{a,b\}$ е линеарно независно множество од вектори тогаш $Q_1 = \{x \mid a^Tx > 0\}$ и $Q_2 = \{x \mid b^Tx > 0\}$ имаат непразен пресек.

<u>Заклучок 2</u>. $Q_1 \cap Q_2 \neq \emptyset$ за отворените полупростори $Q_1 = \{x \in R^n \mid h_{kj}^T x > 0\}$ и $Q_2 = \{x \in R^n \mid h_{jk}^T x > 0\}$ определени со векторите

$$\begin{aligned} h_{kj} &= [0 \ \dots \ 0 \ \ h_{kk} \ 0 \ \dots \ 0 \ -h_{kj} \ 0 \ \dots \ 0]^T \\ h_{jk} &= [0 \ \dots \ 0 \ -h_{jk} \ 0 \ \dots \ 0 \ \ h_{jj} \ 0 \ \dots \ 0]^T \end{aligned} \qquad (3.1)$$

за произволно избрани $k,j \in N$, $k \neq j$, каде што $h_{kj}=h_{jk}=s_k^T s_j$ за обучувачкото множество $S = \{s_j \in R^m \mid j \in N\}$.

Сега, враќајќи се на неравенството (2.7) можеме да заклучиме дека секој вектор на учење p за препознавање на даден лик s_k се содржи во десниот отворен полупростор $Q_1 = \{x \in R^n \mid h_{kj}^T x > 0\}$, каде што векторот h_{kj} е одреден како во Заклучокот 2., и во позитивниот ортант $\{x \in R^n \mid x > 0\}$, т.е. во нивниот пресек. Затоа е од интерес утврдувањето на

<u>Лема 3.2</u>. За обучувачкото множество ненегативни целобројни вектори $S = \{s_j \in R^m \mid j \in N\}$ и n-векторот h_{kj}, $k,j \in N$, $k \neq j$, секое множество

$$M_{kj} = \{x \in R^n \mid h_{kj}^T x > 0, \ x > 0\}, \ k \neq j \qquad (3.2)$$

е непразно множество.

Доказ. Во случајот $h_{kj} = 0$ секој позитивен вектор, $x > 0$ се содржи во M_{kj} и $M_{kj} \neq \emptyset$. Ако пак $h_{kj} > 0$, тогаш, согласно Шварцовото неравенство, е точно дека:

$$h_{kj} < \sqrt{h_{kk}} . \sqrt{h_{jj}}$$

и оттука следува, дека, на пример, векторот $\bar{x}=[\bar{x}_i]_{nx1}$, каде што $\bar{x} = \begin{cases} \sqrt{h_{jj}}/\sqrt{h_{kk}} & \text{за} \quad i=k \\ 1 & \text{за} \quad i\neq k \end{cases}$ припаѓа на M_{kj}, што значи дека $M_{kj} \neq \emptyset$.

Лема 3.3. За M_{kj} и M_{jk}, $k\neq j$, дефинирани како во Лема 3.2., пресекот $D_{kj} = M_{kj} \cap M_{jk}$ е непразно множество.

Доказ. Во случајот $h_{kj} = h_{jk} = 0$, секој $x > 0$ се содржи во D_{kj}. Ако $h_{kj} = h_{jk} > 0$, тогаш условите $h_{kj}^T x > 0$, $h_{jk}^T x > 0$ се сведуваат на

$$h_{kk}x_k - h_{kj}x_j > 0$$
$$-h_{jk}x_k + h_{jj}x_j > 0$$

или, еквивалентно,

$$0 < \frac{h_{jk}}{h_{jj}}x_k < x_j < \frac{h_{kk}}{h_{kj}}x_k$$

Бидејќи $\frac{h_{jk}}{h_{jj}} < \frac{h_{kk}}{h_{kj}}$, може да се избере доволно мал број $\varepsilon > 0$ така што

$$\frac{h_{jk}}{h_{jj}} < \frac{h_{kk}}{h_{kj}} - \varepsilon < \frac{h_{kk}}{h_{kj}}$$

Тогаш, векторот $\bar{x} = [\bar{x}_i]_{nx1}$, каде што

$$\bar{x}_i = \begin{cases} h_{kk}/h_{kj}-\varepsilon & \text{за } i = j \\ 1 & \text{за } i \neq j \end{cases}$$

му припаѓа на D_{kj}.

Заклучок 1. $M_{kj} = \{x\in R^n \mid h_{kj}^T x > 0,\ x > 0\}$ содржи целоброен позитивен вектор.

Случајот $h_{kj} = 0$ е тривијален и затоа нека $h_{kj} > 0$. Тогаш условот $h_{kj}^T x > 0$ се сведува на $(h_{kk}/h_{kj})x_k > x_j$.

Ако $\frac{h_{kk}}{h_{kj}} \geq 1$, тогаш, на пример $\bar{x} = [\bar{x}_i]_{nx1}$, каде што $\bar{x}_i = \begin{cases} 2 \text{ за } i=k \\ 1 \text{ за } i\neq k \end{cases}$ е целоброен и припаѓа на M_{kj}.

Ако $0 < \frac{h_{kk}}{h_{kj}} < 1$, тогаш $\frac{h_{kj}}{h_{kk}} > 1$ и од условот $x_k > \frac{h_{kj}}{h_{kk}} x_j$ се гледа дека векторот $\bar{x} = [\bar{x}_i]_{nx1}$ каде што

$$\bar{x}_i = \begin{cases} [\sqrt{h_{kj}}/\sqrt{h_{kk}}] + 1 & \text{за } i = k \\ 1 & \text{за } i \neq k \end{cases}$$

му припаѓа на M_{kj}.

Согласно (2.7) и (3.2), за препознавање на ликот s_k во однос на секој лик $s_j \in S$, $j\neq k$ потребно е p да припаѓа на секое од множествата M_{kj}, $j\in N$, $j\neq k$ т.е. на нивниот пресек

$$M_k = \bigcap_{j\in N\setminus\{k\}} M_{kj}$$

Значи, p треба да припаѓа на множеството

$$M_k = \{x\in R^n \mid h_{kj}^T x > 0,\ x>0, j\in N\setminus\{k\}\} \tag{3.3}$$

Невронската мрежа за да го препознава секој лик $s_k \in S$, $k\in N$, потребно е да постои цел позитивен вектор $p=p^*$ од пресекот M на множествата M_k, $k\in N$, дефинирани во (3.3), т.е. на

$$M = \{x\in R^n \mid h_{kj}^T x > 0,\ x > 0,\ k,j\in N,\ k\neq j\}, \tag{3.4}$$

кое што множество ја претставува <u>областа на векторите на учење</u> p во R^n.

Прво ќе покажеме дека областа M на векторите на учење дефинирана со (3.4) не е празно множество.

Навистина, за единечните вектори e_i, $i\in N$ и позитивните скалари $\lambda_i = \frac{1}{\sqrt{h_{ii}}}$, $i\in N$, векторот $\bar{x} = \sum_{i=1}^{n} \lambda_i e_i$ очигледно е позитивен и притоа, за секои $k\neq j$, $k,j\in N$, е точно дека

$$h_{kj}^T \bar{x} = \sum_{i=1}^{n} \frac{1}{\sqrt{h_{ii}}} h_{kj}^T e_i = \frac{1}{\sqrt{h_{kk}}} h_{kk} - \frac{1}{\sqrt{h_{jj}}} h_{kj} =$$

$$= \frac{\sqrt{h_{kk}}\sqrt{h_{jj}} - h_{kj}}{\sqrt{h_{jj}}} > 0$$

што значи дека $\bar{x}\in M$.

Поради важноста на областа на векторите на учење M, нејзините основни особини ги резимираме во вид на теорема.

<u>Теорема 3.1</u>. Областа на векторите на учење

$$M = \{x \in R^n \mid h_{kj}^T x > 0,\ x > 0,\ k,j \in N,\ k \neq j\}$$

(i) е непразно, отворено, конвексно множество;

(ii) има својство на конус (заедно со секоја своја точка $\bar{x}$ ја содржи и полуправата $\{x = \lambda\bar{x} \mid \lambda > 0\}$);

(iii) е ограничено од долу и неограничено од горе;

(iv) содржи целоброен (позитивен) вектор.

<u>Доказ</u>. Особините (i), (ii), (iii) се очигледни и затоа ќе се задржиме на докажувањето на (iv). Бидејќи M е непразно и секоја негова точка $\bar{x}$ е внатрешна, постои δ околина $S(\bar{x},\delta) \subset M$. Нека $L = \{\ell = \lambda\bar{x} \mid \lambda > 0\}$ и $u^+ = \{u = y + \ell \mid y \in S,\ \ell \in L\}$, тогаш согласно теоремата на Кронекер [1], u^+ содржи целоброен вектор, кој по конструкција е позитивен.

4. Дефинирање на задачата на целобројното програмирање

Како што беше кажано, по аналогија на човекот, кој за да може да препознае некој лик мора да го видел барем еднаш, и невронската мрежа за да може да препознае некој лик од обучувачкото множество, мора тој лик да ѝ се претстави барем еднаш. Во наставниот простор R^n за компоненти на векторот на учење се земаат броевите на учење на мрежата за ликовите од обучувачкото множество. Бидејќи секој лик од обучувачкото множество мора да ѝ се претстави на мрежата барем еднаш, следи дека секоја компонента на векторот на учење е целобројна и најмалку еднаква на 1. Односно, векторите на учење се позитивни вектори од областа на векторите на учење M т.е. $x = p$, каде што $p = [p_i]_{nx1}$, $p_i \geq 1$ и p_i - целобројни.

Според ова, најмал вектор на учење може да биде векторот

$$e = [1\ 1\ \ldots\ 1]^T$$

а секој друг вектор на учење ги исполнува условите:

- $p \geq e$ и

- p целоброен (и позитивен).

За произволно избрани, $k,j \in N$, $k \neq j$, векторот h_{kj} има целобројни компоненти и затоа за секој целоброен вектор x, скаларниот производ $h_{kj}^T x$ е целоброен. Тогаш условот $h_{kj}^T x > 0$ означува дека $h_{kj}^T x$ е цел позитивен број, така што овој услов за $x \geq e$ и x-целоброен е еквивалентен со условот $h_{kj}^T x \geq 1$. Значи од множеството M, како област на векторите на учење, можеме да преминеме кон неговото вистинско подмножество:

$$D = \{x \in \mathbb{R}^n \mid h_{kj}^T \geq 1,\ x \geq e,\ k,j \in N,\ k \neq j,\ x\text{-целоброен}\}, \qquad (4.1)$$

кое претставува конвексно многустрано множество од $\mathbb{R}^n$, ограничено од долу.

Без да се изгуби ниеден вектор на учење p, сите понатамошни разгледувања за векторите на учење ќе ги вршиме во множеството D. Согласно Теорема 3.1., D е непразно и заедно со секој целоброен вектор p ги содржи и векторите λp, за $\forall \lambda > 0$ и λ-целоброен, што значи дека D е неограничено од горе.

Изборот на правилото на учење ни гарантира постоење на конечен вектор на учење p^* (со конечни позитивни целобројни компоненти) за препознавање на секој лик од S.

Нека е $p = [p_i]_{n\times1}$ (на некој начин) најден вектор на учење т.е. $p \in D$. Понатамошното обучување на мрежата е непотребно, бидејќи со p е дефинирана меморијата на невронската мрежа W, која го препознава секој лик од S. Притоа, заедно со p, ваков вектор е и λp за секој $\lambda > 0$ и λ-целоброен. Оттука следи дека најрационален ќе биде изборот на векторот на учење $p^* = [p_i^*]_{n\times1}$ добиен со најмал број на учења на мрежата за ликовите од S т.е. со најмал сумарен збир на компонентите на p,

$$z^* = p_1^* + \dots + p_n^*$$

или пократко

$$z^* = e^T p^* \leq e^T p \text{ за секој друг вектор на учење } p \in D. \qquad (4.2)$$

За произволен вектор $x \in \mathbb{R}^n$ пресликувањето $z\colon \mathbb{R}^n \to \mathbb{R}$ дефинирано со:

$$z = e^T x$$

претставува линеарен функционал и условот (4.2) означува дека $z^* = e^T p^*$ е минимална вредност на функционалот z, која се достигнува на множеството D, (4.1).

Тоа значи дека проблемот на обучување на невронската мрежа, за наоѓање на вектор на учење што го исполнува условот (4.2), може да се формулира како задача на целобројното линеарно програмирање со функција на целта $z = e^T x$ и допустлива област D, т.е.

$$\min\{z=e^T x \mid h_{kj}^T x \geq 1,\ k,j \in N,\ k \neq j,\ x \geq e,\ x\text{-целоброен}\} \tag{4.3}$$

Оваа задача може да се запише и во следнава компактна форма

$$\min\{z=e^T x \mid \begin{bmatrix} A \\ E \end{bmatrix} x \geq \begin{bmatrix} f \\ e \end{bmatrix},\ x\text{-целоброен}\}$$

каде што E е единечната n-матрица, e е n-вектор, а f е (n-1)n-вектор со сите компоненти еднакви на 1 и A е (n-1)nxn-матрица со следнава структура

$$A = \begin{bmatrix} A_1 \\ \cdots \\ A_k \\ \cdots \\ A_n \end{bmatrix}$$

каде што блокот $A_k = [a_i^{(k)}]_{(n-1)xn}$ има облик

$$A_k = \begin{bmatrix} -h_{k1} & 0 & \cdots & 0 & h_{kk} & 0 & \cdots & 0 \\ 0 & -h_{k2} & & 0 & h_{kk} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -h_{k,k-1} & h_{kk} & 0 & \cdots & . \\ 0 & 0 & \cdots & 0 & h_{kk} & -h_{k,k+1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & h_{kk} & 0 & \cdots & -h_{kn} \end{bmatrix}$$

Заклучок

Обучувањето на невронската мрежа за препознавање ликови може да се сведе на наоѓање целоброен позитивен вектор во векторскиот простор R^n. Во позитивниот ортант на R^n постои отворена конвексна област на векторите на учење, која е непразно ограничено од долу и неограничено од горе множество со својство на конус. Во оваа област се наоѓа вистинско подмножество, кое ги содржи векторите на учење за обучувачкото множество. За наоѓање на оптималниот вектор на учење, се дефинира задача на минимизација од целобројното линеарно програмирање.

Л И Т Е Р А Т У Р А

[1] Белоусов, Е.: Введение в выпуклый анализ и целочисленное программирование; Москва, 1977

[2] Bozinovski, S.: Teaching Space: A representation concept for adaptive pattern classification; COINS Tehnical Report 81-88, University of Massachusetts at Amherst, 1981

[3] Bozinovski, S.: A Representation Theorem for Linear Pattern Classifier Training; IEEE Transactions on Systems, Men and Cibernetics, Vol.SMC-15, 159-161, No.1,January/February, 1985

[4] Simpson, P.: Artificial neural systems: Foundations, Paradigms, Applications, and Implementations; Pergamon Press, San Diego, 1990

[5] Непосредна комуникација на авторите со проф. д-р Д.Карчицка

THE NEURAL NETWORK LEARNING AS A SPECIAL INTEGER LP-PROBLEM

Ǵorǵi Jovančevski & Stevo Božinovski*

S u m m a r y

It is shown that the Neural Network Training can be defined as an Integer Linear Program

$$\min\{z=e^{T}x \mid A_k x \geq \bar{e},\ k=1,\dots,n,\ x \geq e \text{ and integer}\}$$

where

$$A_k = \left[a_{ij}^{(k)}\right]_{(n-1)xn}, \quad a_{ij}^{(k)} = \begin{cases} h_{kk}, & i=1,\dots,n-1,\ j=k \\ -h_{kj}, & i=1,\dots,k-1,\ j=i \\ -h_{kj}, & i=k,\dots,n-1,\ j=i+1 \\ 0, & \text{otherwise} \end{cases}$$

e is n-vector and $\bar{e}$ is n-1-vector with all the components equal 1. The coefficients h_{kj}, k,j=1,...,n are non negative reals given by the inner products of the neural network's learning set of patterns. The set of the feasible solutions is nonempty and bounded from below.

Prirodno-matematički fakultet
p.f. 162
91000 Skopje, Macedonia

*Elektrotehnički fakultet
p.f. 574
91000 Skopje, Macedonia