Математички Билтен
22 (XLVIII)
1998 (61–68)
Скопје, Македонија

ISSN 0351-336X

EXPLICIT SOLUTION OF THE LP-MODEL OF THE NEURAL NETWORK LEARNING PROBLEM

Dimitra Karčicka and Gjorgji Jovančevski

Abstract

In [2] the neutral network learning problem was formulated as the following LP-problem:

Find $x_j \geq 0,\ j = 1,\dots,n+1$ which satisfy the n^2+1 constraints:

$$\begin{aligned} x_i - a_{si}x_s + h_i x_{n+1} &\geq 0\,, \qquad i \neq s\,, \quad i,s = 1,\dots,n \\ x_i \qquad - \qquad x_{n+1} &\geq 0\,, \qquad i = 1,\dots,n \\ x_1 + x_2 + \dots + x_n &= 1 \end{aligned}$$

and maximize the linear form $z = x_{n+1}$;
all the a_{si}, h_i, $0 \leq a_{si} < h_i \leq 1$ are assumed to be known. Then, the feasible bases to the dual problem in standard form were discussed.

Now, we consider directly the above stated LP-problem and characterize the extreme points of the set of feasible solutions.

Let D be the set of feasible solutions of the considered LP-problem, i.e. the set of vectors $\boldsymbol{x} = [x_j]$ from the Euclidean space E^{n+1}, whose components satisfy the conditions:

$$x_i - a_{si}x_s - h_i x_{n+1} \geq 0\,, \qquad i \neq s\,, \quad i,s = 1,\dots,n \tag{1}$$

$$x_i \qquad - x_{n+1} \geq 0\,, \qquad i = 1,\dots,n \tag{2}$$

$$x_1 + x_2 + \dots + x_n = 1 \tag{3}$$

$$x_j \geq 0, \qquad j = 1,\dots,n+1 \tag{4}$$

for given a_{si}, h_i, $0 \leq a_{si} < h_i \leq 1$, $s \neq i$, $i, s = 1, \ldots, n$, and maximize $z = x_{n+1}$.

Obviously, $\boldsymbol{x} = \frac{1}{n}\,[1 \ldots 1\ 0]^{\mathrm{T}} \in D$ (T means transposition).

Since $D \neq \emptyset$ and $0 \leq x_{n+1} \leq x_i \leq 1$ for every $x \in D$ by (2) and (3), it follows that D is a convex polyhedron. Then, D has a finite number of exterme points and, by the fundamental theorem of linear programming, at least one exterme point of D will be an optimal solution. Moreover, any extreme point of D can be found as a solution to a regular system of $n+1$ tight constraints from (1)–(4). Obviously, equation (3) is includes in every such system. The special structure of these systems, and the sparcity of their matrices enable the construction of the extreme points and the selection of one, which is better than that previously considered.

The vector $\boldsymbol{x}^0 = [x_j^0]$, where $x_j^0 = 1/n$, $j = 1, \ldots, n+1$, is a good point to start the considerations, because it is a positive solution to

$$\begin{aligned} x_i \quad - \quad x_{n+1} &= 0\,, \qquad i = 1, \ldots, n \\ x_1 + x_2 + \cdots + x_n &= 1 \end{aligned}$$

which is one of the regular systems of $n+1$ tight constraints, $B\boldsymbol{x} = \boldsymbol{e}_{n+1}$, where $\boldsymbol{e}_{n+1}$ is the $(n+1)$-th unity vector in any case.

It is easy to see that $\boldsymbol{x}^0$ is a feasible solution, and moreover it is an optimal solution, if

$$1 - a_{si} - h_i \geq 0, \qquad s \neq i, \quad i, s = 1, \ldots, n\,.$$

Otherwise, first of all we can pay attention to regular systems of $n+1$ tight constraints, whose matrices have a maximal number $(n+1)$ of nonzero elements in one column. One type of these matrices can be partitioned as follows:

$$B = \begin{bmatrix} A_{\mathrm{I}_k \mathrm{I}_k} & A_{\mathrm{I}_k \mathrm{J}_k} \\ A_{\mathrm{J}_k \mathrm{I}_k} & A_{\mathrm{J}_k \mathrm{J}_k} \end{bmatrix}$$

where $A_{\mathrm{I}_k \mathrm{I}_k}$ is diagonal matrix of $(n-1)$ order, whose diagonal is

$$\mathrm{diag}(-a_{1k}, -a_{2k}, \ldots, -a_{k-1,k}, -a_{k+1,k}, \ldots, -a_{nk})\,,$$

$$A_{\mathrm{I}_k \mathrm{J}_k} = \left[\boldsymbol{e}^{(n-1)} \ - h_k \boldsymbol{e}^{(n-1)}\right], \ A_{\mathrm{J}_k \mathrm{I}_k} = \begin{bmatrix} (\boldsymbol{e}_l)^{\mathrm{T}} \\ (\boldsymbol{e}^{(n-1)})^{\mathrm{T}} \end{bmatrix}, \ A_{\mathrm{J}_k \mathrm{J}_k} = \begin{bmatrix} -a_{kl} & -h_l \\ 1 & 0 \end{bmatrix}$$

for given k and $l \neq k$; $\boldsymbol{e}^{(n-1)}$ denotes the $(n-1)$ vector whose components are all 1; $\boldsymbol{e}_l$ denotes the l-th unity $(n-1)$ vector;

$$I_k = \{1, \ldots, k-1, k+1, \ldots, n\}, \ J_k = \{k, n+1\}\,,$$

$$\underline{J}_k = \{k \in J_l, 1+n^2\}, \quad J_l = \{1,\ldots,l-1,l+1,\ldots,n\}.$$

The inverse B^{-1}, partitioned in the same way as B, is

$$B^{-1} = \frac{1}{\alpha}\begin{bmatrix} B_{\mathrm{I}_k\mathrm{I}_k} & B_{\mathrm{I}_k\mathrm{J}_k} \\ B_{\underline{\mathrm{J}}_k\mathrm{I}_k} & B_{\underline{\mathrm{J}}_k\mathrm{J}k} \end{bmatrix}$$

where

$$\alpha = \frac{\beta_l}{a_{lk}} + \beta_k\gamma, \qquad \beta_l = h_l a_{lk} + h_k, \qquad \beta_k = h_k a_{kl} + h_l, \qquad \gamma = \sum_{\substack{s=1\\ s\neq k}}^{n} \frac{1}{a_{sk}},$$

$$\delta = 1 - a_{lk}a_{kl},$$

$$(B_{\mathrm{I}_k\mathrm{I}_k})_{sj} = \begin{cases} (\beta_k - \alpha a_{sk})/a_{sk}^2, & s = j \neq l \\ (\beta_k + h_k - \alpha a_{lk})/a_{lk}^2, & s = j = l \\ \beta_k/(a_{sk}a_{jk}), & s \neq j \neq l \\ (\beta_k + h_k)/(a_{sk}a_{lk}), & s \neq j = l \end{cases},$$

$$(B_{\mathrm{I}_k\mathrm{J}_k})_{sj} = \begin{cases} h_k/a_{sk}, & s \in I_k, j = k \\ \beta_k/a_{sk}, & s \in I_k, j = n+1 \end{cases}$$

$$(B_{\underline{\mathrm{J}}_k\mathrm{I}_k})_{ij} = \begin{cases} \beta_l/(a_{jk}a_{lk}), & i = 1,\ j \in J_k - \{l\} \\ \beta_l/a_{lk}^2 - h_k\gamma/a_{lk}, & i = 1,\ j = l \\ \delta/(a_{jk}a_{lk}), & i = 2,\ j \in J_k - \{l\} \\ \delta/a_{lk}^2 - (1+\gamma)/a_{lk}, & i = 2,\ j = l \end{cases},$$

$$(B_{\underline{\mathrm{J}}_k\mathrm{J}_k})_{ij} = \begin{cases} -h_k\gamma, & i = j = k \\ \beta_l/a_{lk}, & i = k, \quad j = n+1 \\ -(1+\gamma), & i = n+1, \quad j = k \\ \delta/a_{lk}, & i = j = n+1. \end{cases}$$

Practically, to find the solution $\overline{\boldsymbol{x}} = B^{-1}\mathbf{e}_{n+1}$ and to verify its feasibility, we need only the last column of B^{-1}, which means the second columns of $B_{\mathrm{I}_k\mathrm{J}_k}$ and $B_{\underline{\mathrm{J}}_k\mathrm{J}_k}$.

Thus,

$$\overline{\boldsymbol{x}} = \frac{1}{\alpha}[\overline{x}_s], \qquad \overline{x}_s = \begin{cases} \beta_k/a_{sk}, & s \in I_k \\ \beta_l/a_{lk}, & s = k \\ \delta/a_{lk}, & s = n+1 \end{cases}$$

and $\overline{\boldsymbol{x}}$ is a feasible solution if

$$\frac{\delta}{a_{lk}} \leq \min\left\{\frac{\beta_l}{a_{lk}}, \min_{s\in I_k}\left\{\frac{\beta_k}{a_{sk}}\right\}\right\}$$

and

$$\frac{\beta_k}{a_{ik}} - \frac{a_{si}\beta_k}{a_{sk}} - \frac{h_i\delta}{a_{lk}} \geq 0,$$

$$i \in I_l - \{k\},\ \ i \in I_r = \{1, \ldots, r-1, r+1, \ldots, n\}\ \ r \neq k, l;$$

$\overline{z} = \delta/a_{lk}$ is the corresponding value of the objective function.

Once we have the feasible solutions $\overline{\boldsymbol{x}}$, then in order to find a new system $B'\boldsymbol{x} = \boldsymbol{e}_{n+1}$, we look for the next candidates k' and $l' \neq k'$, such that

$$\delta' = 1 - a_{l'k'}\, a_{k'l'}\,, \qquad \beta_{l'} = h_{l'}\, a_{l'k'} + h_{k'}, \quad \beta_{k'} = h_{k'}\, a_{k'l'} + h_{l'}\,,$$

$$\gamma' = \sum_{\substack{s=1\\ s\neq k}}^{n} \frac{1}{a_{sk'}}\,, \qquad \alpha' = \beta_{l'}/a_{k'\gamma'}$$

and

$$\alpha < \alpha' \tag{5}$$

If there are such k' and l', then we compute the $(n+1)$-th column of $(B')^{-1}$ and we verify if it represents a feasible solution $\boldsymbol{x}'$. The improvement of the objective function is insured by (5). If there are not such k' and l', then we can compute all the other elements of B^{-1} and check the optimality of $\overline{\boldsymbol{x}}$, applying for example, the complementarity theorem.

A little different type of regular system of $n+1$ equations is the one whose matrix can be partitioned as follows:

$$\hat{B} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where $A_{11} = I^{(n-1)}$ is the identity matrix of order $(n-1)$,

$$(A_{12})_{sj} = \begin{cases} -a_{ks}\,, & s \neq k,\ j = 1 \\ -h_s\,, & s \neq k,\ j = 2 \end{cases}$$

$$(A_{21})_{ij} = \begin{cases} -a_{lk}\,, & i = 1,\ j = l \\ 0\,, & i = 1,\ j \neq l \\ 1 & i = 2,\ j \neq k\,, \end{cases} \qquad A_{22} = \begin{bmatrix} 1 & -h_k \\ 1 & 0 \end{bmatrix}$$

for some k and $l \neq k$. Again, the elements of the inverse

$$\hat{B}^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

can be easily computed.

In particular, the elements of the last columns of B_{12} and B_{22}, which means the solutions of the considered system is $\hat{\boldsymbol{x}} = \dfrac{1}{\eta}[\hat{x}_s]$, where

$$\eta = (1 - a_{lk}a_{kl}) \sum_{s \neq k} h_s + (h_k + a_{lk}h_l)\Big(1 + \sum_{s \neq k} a_{ks}\Big).$$

$$\hat{x}_s = a_{ks}(h_k + a_{lk}h_l) + h_s(1 - a_{lk}a_{kl}), \quad s \neq k, l$$
$$\hat{x}_l = h_l + a_{kl}h_k, \quad \hat{x}_k = h_k + a_{lk}h_l, \quad \hat{x}_{n+1} = 1 - a_{kl}a_{lk}.$$

Then, we continue the discussion of feasibility and optimality of $\hat{\boldsymbol{x}}$ as in the previous case. When there is no optimal solution of this type, we can turn our attention to regular systems of $(n+1)$ tight constraints, whose matrices have the structure (may be after rearrangement)

$$B = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where

$$(A_{11})_{ij} = \begin{cases} 1, & i = j = 1, \ldots, n-1 \\ -a_{i+1,i}, & i = 1, \ldots, n-2 \\ 0, & \text{in any other case,} \end{cases}$$

$$(A_{12})_{ij} = \begin{cases} 0, & i = 1, \ldots, n-2, \quad j = 1 \\ -a_{n,n-1}, & i = n-1, \quad j = 1 \\ -h_i, & i = 1, \ldots, n-1, \quad j = 2 \end{cases}$$

$$(A_{21})_{ij} = \begin{cases} 0, & i = 1, \quad j = 1, \ldots, l-1, l+1, \ldots, n-1, \\ -a_{ln}, & i = 1, \quad j = l, \\ 1, & i = 2, \quad j = 1, \ldots, n-1, \end{cases}$$

$$A_{22} = \begin{bmatrix} 1 & -h_n \\ 1 & 0 \end{bmatrix}$$

for given l, $1 \le l \le n$. Since A_{11}^{-1} is found,

$$(A_{11}^{-1}) = \begin{cases} 1, & i = j \\ 0, & i > j \\ \prod_{s=i}^{j-1} a_{s+1,s}, & i < j \end{cases}$$

then the blocks of the inverse B^{-1} can be easily computed, partitioned in the same way as B. So, we find

$$B^{-1} = \frac{1}{\beta} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

where

$$\beta = \Big(1 - a_{ln} \prod_{i=l}^{n-1} a_{i+1,i}\Big)\Big(\sum_{i=1}^{n-1} h_i + \sum_{k=2}^{n-1}\sum_{s=k}^{n-1} h_s \prod_{i=k-1}^{s-1} a_{i+1,i}\Big) + \\ + \Big(1 + \sum_{s=1}^{n-1}\prod_{i=s}^{n-1} a_{i+1,i}\Big)\Big(h_n + a_{ln}\Big)\Big(h_l + \sum_{s=l}^{n-2} h_{s+1}\prod_{i=l}^{s} a_{i+1,i}\Big)$$

$$B_{22} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix},$$

$$\beta_{11} = \sum_{i=1}^{n-1} h_i + \sum_{k=2}^{n-1}\sum_{s=k}^{n-1} h_s \prod_{i=k-1}^{s-1} a_{i+1,i},$$

$$\beta_{12} = h_n + a_{ln}\Big(h_l + \sum_{s=l}^{n-2} h_{s+1}\prod_{i=l}^{s} a_{i+1,i}\Big)$$

$$\beta_{21} = -\Big(1 + \sum_{s=1}^{n-1}\prod_{i=s}^{n-1} a_{i+1,i}\Big), \qquad \beta_{22} = 1 - a_{ln}\prod_{i=l}^{n-1} a_{i+1,i}$$

Then, for the elements of

$$\begin{aligned} B_{12} &= -A_{11}^{-1} A_{12} B_{22}, \\ B_{21} &= -B_{22} A_{21} A_{11}^{-1}, \\ B_{11} &= A_{11}^{-1}(\beta I^{(n-1)} - A_{12} B_{21}) \end{aligned}$$

formulae for computations can also be found.

Actually, we need only the value β_{22}/β in order to conclude that $B^{-1}\boldsymbol{e}_{n+1}$ may be a better candidate for extreme point. Then, we can compute the second column of B_{12}, whose elements are

$$x_i^* = \frac{1}{\beta}\Big(\beta_{12}\prod_{s=i}^{n-1} a_{s+1,s} + \beta_{22}\Big(h_i + \prod_{s=i+1}^{n-1} h_s \prod_{k=i}^{s-1} a_{k+1,k}\Big)\Big), \quad i = 1, \ldots, n-1 .$$

Since $x_n^* = \beta_{12}/\beta$ and $x_{n+1}^* = \beta_{22}/\beta$ are arleady known, we have the solution $\boldsymbol{x}^*$ to the considered system, and we can check if it satisfies the remainder of constraints, and if there are better candidates for feasible solutions of this type.

The other possibilities for extreme points are the solutions to systems whose matrices have almost a block-diagonal form, with blocks as the matrices considered above.

This process of solving the considered LP-problem seems effective, because the formulae and the computations are particularly simple.

References

[1] Jovančevski Gj., Karčicka D., Božinovski S.: *Neural network training as a special integer linear programming problem*, LIRA'95, Novi Sad 26-30 September, FR Yugoslavia.

[2] Karčicka D., Jovančevski Gj.: *LP-Model of the neural network training problem in the teaching space*, LIRA'95, Novi Sad 26-30 September, FR Yugoslavia.

[3] Karmanov V.G.: *Mathematical programming*, Nauka, Moscow, 1980.

[4] Schrijver A.: *Theory of linear and integer programming*, John Wiley & Sons, 1986.

ЕКСПЛИЦИТНО РЕШЕНИЕ НА LP–МОДЕЛ НА ПРОБЛЕМОТ НА ОБУЧУВАЊЕ НЕВРОНСКА МРЕЖА

Димитра Карчицка и Ѓорѓи Јованчевски

Р е з и м е

Оптимално решение на ЛП-задачата за обучуване на невронска мрежа за препознавање ликови, може да се најде со пребарување на екстремалните точки на конвексниот полиедар од допустливи решенија, користејќи ја едноставната структура на матрицата на регуларните системи ограничувања од $n+1$-ви ред.

Faculty of Natural Sciences and Mathematics,
Institute of Informatics
St. Cyril and Methodius University,
P.Box 162,
91000 Skopje,
Macedonia

dimitra@robig.pmf.ukim.edu.mk

gjorgji@robig.pmf.ukim.edu.mk