KYBERNETIKA ČÍSLO 6, ROČNÍK 5/1969

The Convergence of a Committee Solution of the Pattern Recognition Problem

SVATOPLUK BLÁHA

The paper deals with the problem of the convergence of the error-correction training procedures with a correction increment for the committee solution of pattern recognition problem. The training procedure converges, if the angles between solution vectors and weight vectors are reduced at every step of training process. It has been proved that the described type of training procedures cannot secure the convergence to the solution of the problem. Only stochastic training can be succesful.

1. INTRODUCTION

A basic pattern recognition problem is the assignement of each of the given set of patterns to one of two disjunct subsets. If no limiting conditions are put on the decomposition of pattern in the space, then this problem cannot be generally solved by using single threshold logic unit. The two-layer parallel network of the threshold logic units – the committee machine – working as piecewise linear system has more chance to be successful [1]. The proof that the solution of this problem exists have been published in [2]. Many experiments were made with committee machine with different training procedures. They were published for instants in [1]. The purpose of this paper will be the solution of the convergence problem of training methods for committee machine.

This paper is the continuation of paper [3] and there is very useful for the reader to get acquainted with it. Many notations and mathematical derivations in this paper are the same as in [3] and it is not suitable to repeat them with the exception of the most important of them.

Let us assume, that there exist two disjunct subsets $\mathscr{Y}_1$ and $\mathscr{Y}_2$ of the patterns and its union

$$\mathscr{Y} = \mathscr{Y}_1 \cup \mathscr{Y}_2 \tag{1.1}$$

is the training set. Each pattern is represented by a D-dimensional vector $\mathbf{Y}_k^\circ$ ($k = 1, 2, \ldots$) and one of its components is identically equal to $+1$ (threshold input) [1, ch. 4.2]. It means that the patterns lie in the $(D - 1)$-dimensional hyperplane which is at distance 1 from the origin. The committee machine consists in the first layer of the odd number – P – of threshold logic units. Each unit is represented by a D-dimensional vector $\mathbf{W}^{(i)}$ ($i = 1, 2, \ldots, P$), the components of which are the weights of distinct inputs. The decision hyperplanes, all passing through the origin and normal to the vector $\mathbf{W}^{(i)}$ separate correctly the patterns of both subsets in a part of the hyperplane. The boundary between subsets is put together from parts of decision hyperplanes so that every pattern is put on the correct side of the majority of the decision hyperplanes. The patterns are classified by the signum of a dot product $\mathbf{Y}_k^\circ \,.\, \mathbf{W}^{(i)}$. (For the linear separable subsets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ is the dot product $\mathbf{Y}_k^\circ \,.\, \mathbf{W}^{(i)}$ positive for all patterns belonging into $\mathcal{Y}_1$ and negative for others). Such a dichotomy problem can be solved using the single threshold logic unit [1,3].

For economy of description the following adjustment of the training set $\mathcal{Y}$ will be used.

Let us denote the adjusted pattern vectors by the symbol $\mathbf{Y}'_k$; then

$$\mathbf{Y}'_k = \mathbf{Y}_k^\circ \quad \text{if} \quad \mathbf{Y}_k^\circ \in \mathcal{Y}_1 \,,$$
$$\mathbf{Y}'_k = -\mathbf{Y}_k^\circ \quad \text{if} \quad \mathbf{Y}_k^\circ \in \mathcal{Y}_2 \,.$$

The advantage of this adjustment is the fact that the patterns of both classes will be correctly classified, if the scalar product $\mathbf{Y}'_k \,.\, \mathbf{W}_k^{(i)}$ is positive. The training procedure can be defined only by one equation e.g., (2.1) or (3.1). More details can be found in [1, Ch. 5.2] or in [3, Ch. 1]. The cases when the patterns are correctly classified will be eliminated for the simplification of the notation; only the reduced training sequence of the pattern vectors and the weight vectors will be considered [1,3]. The members of the adjusted and reduced training sequence will be denoted $\mathbf{Y}_k$ and the members of the reduced weight sequence will be denoted $\mathbf{W}_k^{(i)}$.

The convergence conditions for the training procedure depend in the reciprocal positions of the solution vector $\mathbf{W}_*^{(i)}$, the weight vector $\mathbf{W}_k^{(i)}$ and the pattern vector $\mathbf{Y}_k$. These reciprocal positions will be described by the angles between the vectors. Let us denote for the k-th step

$$(1.2) \qquad \sphericalangle\, \mathbf{W}_*^{(i)}, \mathbf{W}_k^{(i)} = \omega_k^{(i)} \,,$$
$$\sphericalangle\, \mathbf{W}_*^{(i)}, \quad \mathbf{Y}_k = \eta_k^{(i)} \,,$$
$$\sphericalangle\, \mathbf{W}_k^{(i)}, \quad \mathbf{Y}_k = \xi_k^{(i)} \,.$$

The training procedure converges to the solution of the problem if the limits of the sequences of the angles between solution vectors $\mathbf{W}_*^{(i)}$ and the weight vectors

 $\mathbf{W}_k^{(i)}$ are:

(1.3)
$$\lim_{k \to \infty} \omega_k^{(i)} = 0 \,, \quad i = 1, 2, \ldots, P \,,$$

where index k denotes k-th step of training process. It is supposed, of course, that any two or more weight vectors do not converge to the same solution vector. The training process can be represented at each step by the points in angle-space.

The values of angles ω_k, η_k and ξ_k are delimitated and the limits are derived from the geometrical positions of the vector in the space [3]:

(1.4)
$$\omega_k^{(i)} - \eta_k^{(i)} < \xi_k^{(i)} \,,$$
$$\xi_k^{(i)} \leqq \omega_k^{(i)} + \eta_k^{(i)} \quad \text{for} \quad \omega_k^{(i)} + \eta_k^{(i)} \leqq \pi \,,$$
$$\xi_k^{(i)} \leqq 2\pi - \omega_k^{(i)} - \eta_k^{(i)} \quad \text{for} \quad \omega_k^{(i)} + \eta_k^{(i)} > \pi \,.$$

As the cases when the patterns are correctly classified are eliminated, it means that only such situations will be considered when

(1.5)
$$\tfrac{1}{2}\pi < \xi_k^{(i)} < \pi \,.$$

The values of angles $\omega_k^{(i)}$ are evidently

(1.6)
$$0 \leqq \omega_k^{(i)} \leqq \pi \,.$$

The values of angles $\eta_k^{(i)}$ can be

(1.7)
$$0 \leqq \eta_k^{(i)} < \tfrac{1}{2}\pi$$

if the training set is linearly separable and

(1.8)
$$0 \leqq \eta_k^{(i)} \leqq \pi \,.$$

in general case. From the relation (1.5) follows

(1.9)
$$\tfrac{1}{2}\pi < \omega_k^{(i)} + \eta_k^{(i)} < \tfrac{3}{2}\pi \,.$$

The region of the permissible angles ω, η and ω, η, ξ are illustrated on Fig. 3.1 and Fig. 3.2.

2. THE CONVERGENCE CONDITIONS FOR SINGLE THRESHOLD UNIT

Let us assume, that for the correction of the weight vector will be used the rule

(2.1)
$$\mathbf{W}_{k+1} = \mathbf{W}_k + c_k \mathbf{Y}_k \,, \quad k = 1, 2, \ldots \,,$$

where the correction increment c_k is the positive constant. If the classifying problem can be solved by the single threshold unit, than for the values of angle η_k the inequation (1.7) is valid (regions $\mathscr{A}_1, \ldots, \mathscr{A}_4$ in Fig. 3.1). The convergence conditions for error-correction training procedures have been discussed in detail in [3]. The result has been following:

1. The representing point must be approaching the point

$$S_1(\eta = \tfrac{1}{2}\pi; \omega = 0).$$

2. In the neighbourhood of the point S_1 it is necessary that the correction increment is sufficiently small. The training procedures converge with this condition in the regions $\mathscr{A}_1$, $\mathscr{A}_2$, $\mathscr{A}_3$ and in the part of region $\mathscr{A}_4$ where the inequality is valid

$$|\cos \xi_k| < \frac{\cos \eta_k}{|\cos \omega_k|}. \tag{2.2}$$

In the remaining part of the region $\mathscr{A}_4$ the training procedures converge only if c_k is sufficiently great. This condition cannot always be fulfilled. The area

$$|\cos \xi_k| = \frac{\cos \eta_k}{|\cos \omega_k|} \tag{2.3}$$

is on Fig. 3.2 illustrated by the dash lines and there is supposed at first, that the couples of vectors $\mathbf{W}_*^{(i)}$ and $\mathbf{W}_k^{(i)}$ are known for all $i = 1, 2, \ldots, P$.

3. THE CONVERGENCE CONDITIONS FOR COMMITTEE MACHINE

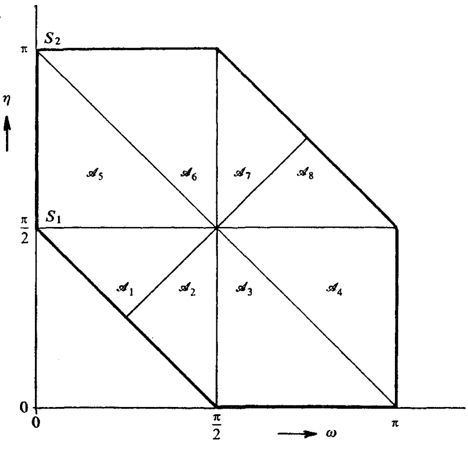
Let us assume, the committee machine has in the first layer P threshold logic units and this number is sufficient for solution of the problem. Each threshold logic unit is characterized by the weight vector $\mathbf{W}_k^{(i)}$, $i = 1, 2, \ldots, P$, $k = 1, 2, \ldots$ If the pattern $\mathbf{Y}_k$ is uncorrectly classified, then is necessary to change the minimum number of weight vectors $\mathbf{W}_k^{(i)}$ to obtain the correct result. The vectors choosen to be changed in certain step k can be selected from the set of uncorrectly classifying vectors in different ways: the magnitude of the weight vectors, the value of the scalar product $\mathbf{W}_k^{(i)} . \mathbf{Y}_k$, the angle $\xi_k^{(i)}$ etc. can each be used. The kinds of selection have an influence upon the convergence of the training process. These methods will be not discused in this paper.

The changes of weight vectors are given by the equations

$$\mathbf{W}_{k+1}^{(i)} = \mathbf{W}_k^{(i)} + c_k^{(i)} \mathbf{Y}_k, \tag{3.1}$$

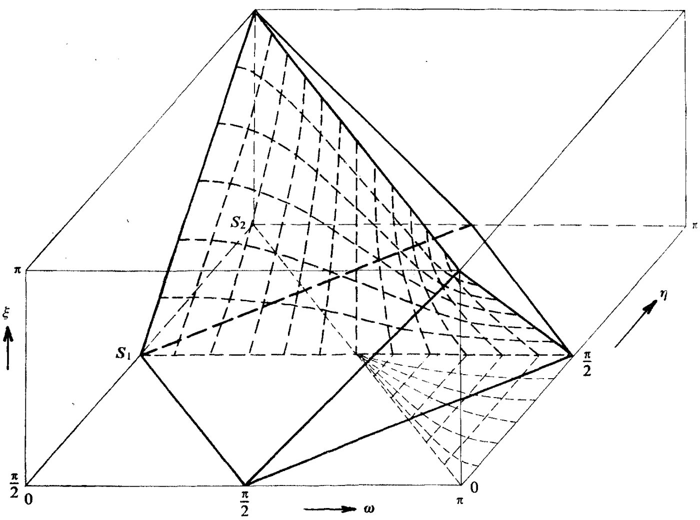
where $c_k^{(i)}$ are positive constants which must be determined. The necessary and the sufficient conditions for the relation (1.3) to be true for arbitrary initial weight vectors and any sequence of pattern vectors are

$$\omega_{k+1}^{(i)} \leqq \omega_k^{(i)}, \quad i = 1, 2, \ldots, P \tag{3.2}$$



S_2
π
η
$\mathscr{A}_5$
$\mathscr{A}_6$
$\mathscr{A}_7$
$\mathscr{A}_8$
S_1
$\frac{\pi}{2}$
$\mathscr{A}_1$
$\mathscr{A}_2$
$\mathscr{A}_3$
$\mathscr{A}_4$
0
0
$\frac{\pi}{2}$
π
ω

Fig. 3.1.



S_2
π
π
ξ
η
S_1
$\frac{\pi}{2}$
$\frac{\pi}{2}$
0
$\frac{\pi}{2}$
π
0
ω

Fig. 3.2.

and the equality is valid, when the vector was not corrected. The problem, if this necessary condition can be always satisfied, will be solved in the bellow. The convergence of the training procedure must be considered in the regions which are illustrated in Fig. 3.1. The convergence condition in regions $\mathscr{A}_1, \ldots, \mathscr{A}_4$ have been discused in detail in [3]. The similar way can be used for regions $\mathscr{A}_5, \ldots, \mathscr{A}_8$ too. The functions $\cos^2 \omega_k$ can be substituted for ω_k in (3.2). The inequations

$$\left(\frac{\mathbf{W}_*^{(i)} \cdot \mathbf{W}_k^{(i)}}{|\mathbf{W}_*^{(i)}|\,|\mathbf{W}_k^{(i)}|}\right)^2 < \left(\frac{\mathbf{W}_*^{(i)} \cdot \mathbf{W}_{k+1}^{(i)}}{|\mathbf{W}_*^{(i)}|\,|\mathbf{W}_{k+1}^{(i)}|}\right)^2 \tag{3.3}$$

is equivalent to the inequation (3.2) if $0 \leqq \omega_k^{(i)} < \frac{1}{2}\pi$. The relation (3.3) is valid conversely if $\frac{1}{2}\pi < \omega_k^{(i)} \leqq \pi$. We shall obtain the different results for distinct regions $\mathscr{A}_5, \ldots, \mathscr{A}_8$. To make the point clear they are given in table (3.4):

		$\cos^2 \omega_k - \cos^2 \eta_k > 0$	$\cos^2 \omega_k - \cos^2 \eta_k < 0$
(3.4)	$0 \leqq \omega_k < \frac{1}{2}\pi$ ($\cos \eta < 0$, $\cos \omega > 0$)	$\mathscr{A}_5$ $c_k < \dfrac{2\lvert\mathbf{W}_k^{(i)}\rvert}{\lvert\mathbf{Y}_k\rvert} \Lambda_k^{(i)}$	$\mathscr{A}_6$ $c_k > \dfrac{2\lvert\mathbf{W}_k^{(i)}\rvert}{\lvert\mathbf{Y}_k\rvert} \Lambda_k^{(i)}$
	$\frac{1}{2}\pi < \omega_k \leqq \pi$ ($\cos \eta < 0$, $\cos \omega < 0$)	$\mathscr{A}_8$ $c_k > \dfrac{2\lvert\mathbf{W}_k^{(i)}\rvert}{\lvert\mathbf{Y}_k\rvert} \Lambda_k^{(i)}$	$\mathscr{A}_7$ $c_k < \dfrac{2\lvert\mathbf{W}_k^{(i)}\rvert}{\lvert\mathbf{Y}_k\rvert} \Lambda_k^{(i)}$

where

$$\Lambda_k^{(i)} = \frac{\cos \omega_k^{(i)} \left(\cos \eta_k^{(i)} + |\cos \xi_k^{(i)}| \cos \omega_k^{(i)}\right)}{\cos^2 \omega_k^{(i)} - \cos^2 \eta_k^{(i)}}\,. \tag{3.5}$$

The training procedures can converge, if the inequations (3.4) are fulfilled for the positive c_k. In the region $\mathscr{A}_5$ is $\cos \eta_k^{(i)} < 0$, $\cos \omega_k^{(i)} > 0$ and the denominator of $\Lambda_k^{(i)}$ is positive. It is necessary for the numerator of $\Lambda_k^{(i)}$ to be positive, too:

$$\cos \omega_k^{(i)}\left(-|\cos \eta_k^{(i)}| + |\cos \xi_k^{(i)}| \cos \omega_k^{(i)}\right) > 0\,. \tag{3.6}$$

To fulfil it, it is necessary that

$$|\cos \xi_k^{(i)}| > \frac{|\cos \eta_k^{(i)}|}{\cos \omega_k^{(i)}}\,. \tag{3.7}$$

The inequations (3.7) and (1.4) define the subregion in $\mathscr{A}_5$, where the training procedures based on the equation (3.1) converge:

$$(3.8)\qquad \arccos \frac{|\cos \eta_k^{(i)}|}{\cos \omega_k^{(i)}} < \xi_k^{(i)} \leqq \eta_k^{(i)} + \omega_k^{(i)} \quad \text{for} \quad \eta_k^{(i)} + \omega_k^{(i)} < \pi ,$$

$$\arccos \frac{|\cos \eta|_k^{(i)}}{\cos \omega_k^{(i)}} < \xi_k^{(i)} \leqq 2\pi - \eta_k^{(i)} - \omega_k^{(i)} \quad \text{for} \quad \eta_k^{(i)} + \omega_k^{(i)} > \pi .$$

The correction increment must be sufficiently small.

In the remaining part of $\mathscr{A}_5$ the numerator of $\Lambda_k^{(i)}$ is negative, i.e. $\Lambda_k^{(i)}$ is negative too and the training procedures can converge to the solution of problem for $c_k^{(i)} < 0$ only, but it conflicts with the supposition that $c_k^{(i)}$ is always positive.

In the region $\mathscr{A}_6$ the denominator of $\Lambda_k^{(i)}$ is negative, because it is evidently

$$(3.9)\qquad \cos \omega_k^{(i)} < |\cos \eta_k^{(i)}| .$$

It means that $|\cos \eta_k^{(i)}|/\cos \omega_k^{(i)} > 1$ and

$$(3.10)\qquad |\cos \xi_k^{(i)}| < \frac{|\cos \eta_k^{(i)}|}{\cos \omega_k^{(i)}}$$

for all $\eta_k^{(i)}$ and $\omega_k^{(i)}$ in $\mathscr{A}_6$. The numerator of $\Lambda_k^{(i)}$ is negative and $\Lambda_k^{(i)}$ positive. By the (3.4) the value of correction increment $c_k^{(i)}$ must be sufficiently great.

The convergence condition (3.3) in $\mathscr{A}_6$, which means that

$$(3.11)\qquad |\cos \omega_{k+1}^{(i)}| \geqq |\cos \omega_k^{(i)}| ,$$

is valid until

$$(3.12)\qquad \omega_{k+1}^{(i)} + \omega_k^{(i)} \leqq \pi .$$

The convergence condition (3.11) should be fulfilled for small changes of the weight vector $\mathbf{W}_k^{(i)}$, because for such changes the Eq. (3.12) will be satisfied. But the convergence conditions (3.4) in the region $\mathscr{A}_6$ is fulfilled only for

$$(3.13)\qquad \omega_{k+1}^{(i)} + \omega_k^{(i)} > \pi$$

and it means, that $\omega_{k+1}^{(i)} > \omega_k^{(i)}$, because in the region $\mathscr{A}_6$ is $\omega_k^{(i)} \leqq \frac{1}{2}\pi$. The conclusion is, that the convergence condition (3.2) cannot be fulfilled in the region $\mathscr{A}_6$ for positive value of correction increment $c_k^{(i)}$.

In the region $\mathscr{A}_7$ is $\cos \omega_k^{(i)} < 0$, $\cos \eta_k^{(i)} < 0$ and $|\cos \omega_k^{(i)}| < |\cos \eta_k^{(i)}|$ so that the denominator of the $\Lambda_k^{(i)}$ is negative for all $\eta_k^{(i)}$ and $\omega_k^{(i)}$ from $\mathscr{A}_7$. The numerator of $\Lambda_k^{(i)}$ is evidently positive. The value of $\Lambda_k^{(i)}$ is negative inside the whole region $\mathscr{A}_7$. By the (3.4) the correction increment would have to be negative for the convergence condition to be fulfilled. In region $\mathscr{A}_7$ the training procedures of this type cannot converge.

In the region $\mathscr{A}_8$ is $|\cos \omega_k^{(i)}| > |\cos \eta_k^{(i)}|$, it means that the denominator of the $\Lambda_k^{(i)}$ is positive. The numerator of $\Lambda_k^{(i)}$ is always positive similarly as in the region $\mathscr{A}_7$. The value of the $\Lambda_k^{(i)}$ is positive, so that the correction increment $c_k^{(i)}$ must be positive and sufficiently great.

The solution of the problem is on the abscissa $\overline{S_1 S_2}$, ($\omega = 0$; $\frac{1}{2}\pi \leqq \eta < \pi$; $\xi = \frac{1}{2}\pi$) (Fig. 3.1). In the neighbourhood of point S_1 the training procedures converge if the correction increment is sufficiently small. Evidently it is impossible to fulfil the different conditions in individual regions $\mathscr{A}_5, \ldots, \mathscr{A}_8$, because we do not know, in which of these regions the representative point is in a certain step.

After correction the representative point moves along the axis ω in Fig. 3.1. This point moves to the left if the convergence condition of the proper region is fulfilled. This case is the convergent step. The point moves to the right in the opposite case. The convergence of the training methods can be secured in the regions $\mathscr{A}_1$, $\mathscr{A}_2$, $\mathscr{A}_3$ and in the parts of the regions $\mathscr{A}_4$ and $\mathscr{A}_5$, if the correction increment is sufficiently small. On the contrary the representative point will move to the right (the angle $\omega^{(i)}$ will increase – the divergent step) if it lies in the remaining part of the region $\mathscr{A}$, because the correction increment must not be negative.

The representative point moves also after input of the new pattern vector and it moves in this case along the axis η. The point can pass from a region, where the training methods do not converge, into a region where they do and vice versa.

It is necessary to consider that the pairs of vectors $\mathbf{W}_*^{(i)}$, $\mathbf{W}_k^{(i)}$ ($i = 1, 2, \ldots, P$) are not fixed. P different representative points correspond to every corrected weight vector $\mathbf{W}_k^{(i)}$; each representative point is determined by three angles between vectors $\mathbf{Y}_k$, $\mathbf{W}_k^{(i)}$, $\mathbf{W}_*^{(i)}$ ($i = 1, 2, \ldots, P$). These points lie in different regions $\mathscr{A}_1, \ldots, \mathscr{A}_8$; it means that convergence conditions can be satisfied for one or several or all representative points, or for none. Considering a particular step, a single weight vector $\mathbf{W}_k^{(i)}$ can converge to a selection of solution vectors $\mathbf{W}_*^{(j)}$ when the convergence conditions are satisfied for their representative points. The weight vector $\mathbf{W}_k^{(i)}$ diverges with regard to all others solution vectors. The distinct weight vectors converge usually to the different solution vectors during the training process.

It is evident that the training process necessarily consists of the convergent and the divergent steps. The training process will converge, if the effect of the convergent steps predominates over the effect of the divergent steps. It is evident, that this ratio depends on the training sequence of the patterns and on the random positions of initial weight vectors $\mathbf{W}_1^{(i)}$ ($i = 1, 2, \ldots, P$). A training sequence can be also considered as a random or quasirandom sequence.

The convergence of the training method is influenced also by the rule for selection of weight vectors which shall be corrected. This influence can be shown in a simple example without specification of the rule. Let us assume, that all weight vectors of the committee machine with the exception of one of them correspond to the solution vectors:

$$W_k^{(j)} \equiv W_*^{(j)} ; \quad j = 1, 2, \ldots, 2r; \quad 2r + 1 = P .$$

(Half of the number of these vectors $W_k^{(j)}$ $(j = 1, 2, \ldots, r)$ classify pattern Y_k correctly and the other half $(j = r + 1, \ldots, 2r)$ do not. So the position of the remaining weight vector $W_k^{(P)}$ is decisive for the classification and it is necessary to change just this vector, to obtain the solution of the problem. But also any of the weight vector $W_k^{(j)}$ $(j = r + 1, \ldots, 2r)$ uncorrectly classifying the pattern Y_k can be selected for the correction. Such a step of the training method is evidently divergent. The selection depends on positions of weight vectors and the pattern vector in the space. It can be considered as a random process regardless of which rule is used.

The sequence of the angles $\{\omega_k^{(i)}\}$ converges to a limit, if and only if an index exists for any small positive number ε, such that for every k_1, $k_2 > k_0$ is $|\omega_{k_1}^{(i)} - \omega_{k_2}^{(i)}| < \varepsilon$. It is clear that the part of a training sequence $\{Y_k\}$ beginning from an arbitrary k can be such, that the condition mentioned above is not satisfied. The conclusion is that: it is not possible to secure the necessary condition for the training procedure of type (3.1) to be convergent for any sequence of pattern vectors Y_k and any initial weight vectors $W_1^{(i)}$.

The solution of the recognition problem can be reached, by luck, as well as by the convergent algorithmus, though the latter is improbable.

4. THE RANDOM SOLUTION

In the previous chapter it has been shown that it is necessary to consider changes in the weight vectors as random changes. Let us investigate the possibility of acquiring a solution for this model.

The accuracy of measurement and adjustement of weight vectors is limited in every practical case. It means that for every weight vector only a finite number of different positions exist. It means that the angle space can be divided only into a finite number of small δ-regions, determined by the small space-angle δ. The solution of the problem is reached, if all the weight vectors are identical with the solution vectors; it is only valid for the case when a single solution of the problem exists. Two vectors are identical (in this supposition) if they lie in the same δ-region.

There exist different probabilities for the transition of each weight vector from its position to another one. The probabilities depend only on the old and the new positions. Such a process is the Markowian process. Let us find, what is the probability that all weight vectors fall simultaneously into solution regions. If we suppose that the number of changes is unbounded then it is evident that

$$\lim_{k \to \infty} p(W_k^{(i)} \in \mathscr{W}^{(i)} ; \quad i = 1, 2, \ldots, P) = 1 , \tag{4.1}$$

where $\mathscr{W}^{(i)}$ are the solution regions around the vectors $W_*^{(i)}$.

(Received January 15th, 1969.)

[1] Nilsson N. J.: Learning Machines, McGraw Hill Book Co., N. Y. 1965.
[2] Ablow C. M., Kaylor D. J.: A Committee Solution of the Pattern Recognition Problem. IEEE Transactions on Information Theory *IT-11* (July 1965), 3, 453–455.
[3] Bláha S.: The Convergence of One Group of Correction Training Procedures. Kybernetika *5* (1962), 2, 136–153.

VÝTAH

Konvergence většinového řešení rozpoznávání obrazů

SVATOPLUK BLÁHA

Základní problém rozpoznávání obrazů, tj. třídění prvků dvou daných disjunktních množin může být vždy řešen učícím se strojem sestaveným z lichého počtu paralelně spojených lineárních prahových jednotek v první vrstvě a jednou majoritní logickou jednotkou v druhé vrstvě (committee machine). Je znám důkaz existence řešení a byly experimentálně ověřeny i trenovací algoritmy pro takový stroj. Článek je věnován konvergenci trenovacích metod s korekčním inkrementem. Konvergence trenovací metody znamená, že se zmenšuje úhel mezi vektory řešení a příslušnými váhovými vektory. Ukázalo se, že žádná z popsaných trenovacích metod nezaručuje konvergenci k řešení problému. Řešení problému je dosaženo na základě zřejmé skutečnosti, že při neomezeném počtu pokusů a při konečném počtu možností nastane žádoucí situace s pravděpodobností rovnou jedné.

Ing. Svatopluk Bláha CSc., Ústav teorie informace a automatizace ČSAV, Vyšehradská 49, Praha 2.