KYBERNETIKA ČÍSLO 5, ROČNÍK 4/1968

Test lineární separability

MILOŠ THOMA

Práce se zabývá možností testování lineární separability při použití prahového logického adaptivního obvodu v modelování procesu učení. Ukazují se možnosti nahrazení učení některou běžnou trenovací metodou testem lineární separability.

1. ÚVOD

Problematika učení a učících se obvodů je již dosti dlouhou dobu oblastí zvýšeného zájmu, v neposlední řadě vzhledem k možnostem aplikace jako učících se regulátorů. Valná většina známých prací, které se zabývají uvedenou aplikací učících se obvodů v regulaci, používá jako základního prvku pro tyto obvody prahového adaptivního logického elementu. Největší zásluha o prosazování tohoto elementu v uvedené problematice se připisuje prof. B. Widrowovi. Ve svých pracech se zabývá problematikou adaptivního logického prahového obvodu a jeho aplikacemi. Pro podobnost tohoto obvodu s lineárním modelem neuronu nazývá jej B. Widrow ADALINE (adaptivní lineární neuron).

2. PRAHOVÝ LOGICKÝ ADAPTIVNÍ OBVOD

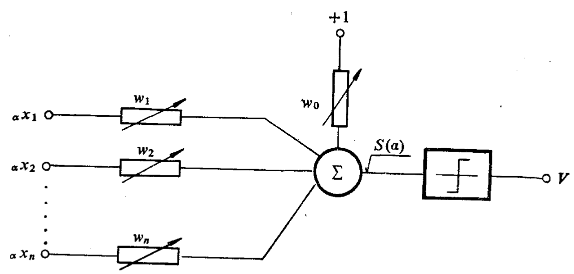
Na obr. 1 je schéma uvedeného obvodu. Vlastní prahový obvod se skládá z n řádných vstupů, z prahového vstupu, ze sumátoru a konečně z výstupního členu (limitéru). Všechny vstupy, včetně prahového, jsou připojeny k sumátoru přes vahové prvky. Velikost vah označme $w_0, w_1, \ldots, w_n$.

Pro funkci obvodu je dále nutné adaptační zařízení, které samo není přímo součástí prahového obvodu a které obstarává přestavování jednotlivých vah. Předpokládejme, že vstupní veličiny mohou nabývat pouze dvou hodnot, a to $+1$ a -1. (Prahový vstup je konstantně připojen na $+1$.)

Nazveme konečně jako obraz každou možnou kombinaci vstupních proměnných $+1$ a -1. Každou jednotlivou vstupní proměnnou označme x_i, $i = 1, \ldots, n$, celý obraz pak symbolem ${}_{\alpha}(x_i)$, kde α je pořadové číslo obrazu.

Mějme dáno s různých obrazů ${}_1(x_i), \dots, {}_s(x_i)$ a to tak, že požadujeme, aby při jejich rozpoznávání byly řazeny obrazy ${}_1(x_i), \dots, {}_r(x_i)$ do prvé a obrazy ${}_{r+1}(x_i), \dots$-$\dots, {}_s(x_i)$ do druhé skupiny. Příslušnost obrazu do prvé skupiny vyznačí prahový obvod v procesu rozpoznávání výstupem (označme V) $+1$, příslušnost do druhé skupiny výstupem -1.

Pokusíme-li se nyní postihnout funkci prahového obvodu, pak jeho činnost musíme rozdělit do procesu učení a procesu rozpoznávání. V procesu učení adaptační zařízení přestavuje váhy v jednotlivých vstupech tak, aby obvod požadovaným způsobem reagoval na jednotlivé obrazy přiváděné na jeho vstupy.



Obr. 1.

V procesu rozpoznávání prahový obvod pak zařazuje přiváděné vstupní obrazy do příslušných tříd podle toho, jak byl „natrénován" v procesu učení.

Všimneme-li si, jakým způsobem obvod provádí zařazení, vidíme, že

(1) $$V = \operatorname{sign} S(\alpha)\,,$$

kde

(2) $$S(\alpha) = w_0 + \sum_{i=1}^{n} {}_\alpha x_i w_i\,,$$

což je součet prahové váhy a skalárního součinu vektoru obrazu ${}_\alpha(x_i)$, $i = 1, \dots, n$ a vektoru vahového $\bar{w} = [w_1, \dots, w_n]$. Pro zpřesnění rozpoznávání se zavádí tzv. pásmo necitlivosti

(3) $$d > 0\,.$$

Potom

(4) $$V = +1\,,$$

jestliže

(5) $$S(\alpha) > d$$

a

$$V = -1 \tag{6}$$

jestliže

$$S(\alpha) < -d\,. \tag{7}$$

Není-li splněna ani jedna z nerovností (5), (7), pak obraz je označen jako neidentifikovatelný.

3. LINEÁRNÍ SEPARABILITA

Pro nalezení váhového vektoru $\bar{w}$, který splňuje zadání úlohy, je známo několik postupů, označovaných jako trénovací metody [1], [2].

Konvergence těchto metod je dokázána za předpokladu, že hledaný váhový vektor existuje. (Existuje-li takový vektor, pak úloha se označuje jako lineárně separabilní. V opačném případě jako lineárně neseparabilní.) Při modelování prahového obvodu a procesu učení (třeba na počítači), nekonverguje-li proces učení, tj. není-li možno nalézt vektor $\bar{w}$, nemůžeme říci, je-li tomu tak proto, že problém je lineárně neseparabilní, nebo že je nekonvergence způsobena nevhodnou volbou pásma necitlivosti. Dále je zřejmé, že řešení – vektor $\bar{w}$ – obvykle není jediné, avšak při použití zmíněných trenovacích metod dostaneme pouze jediné řešení (i když splňující danou úlohu).

Zdá se tedy být opodstatněný požadavek nalezení takového postupu, který by

a) rozhodl o lineární separovatelnosti či neseparovatelnosti dané úlohy,
b) vymezil pokud možno celou množinu řešení.

Zde bude užitečné uchýlit se k prostorové interpretaci celého problému. Obrazy reprezentované n resp. $(n + 1)$ vstupními proměnnými lze považovat za body v n resp. $(n + 1)$-rozměrném prostoru. Dále je dána nadrovina ϱ (s normálovým vektorem $\bar{w}$), která dělí n resp. $(n + 1)$-rozměrný prostor do dvou poloprostorů. Rozpoznávání obrazů dělením do dvou tříd v prostorové interpretaci pak znamená posuzování, ve kterém z obou zmíněných poloprostorů leží uvažovaný bod. Předepíšeme-li pásmo necitlivosti, pak udáváme v nějakém měřítku minimální nutnou vzdálenost bodů od nadroviny. Lze také říci, že zde prahový obvod realizuje lineární rozhodovací funkci. Konkrétní podobou této funkce je zřejmě skalární součin normálového vektoru nadroviny $\bar{w}$ s polohovým vektorem (radiusvektorem) bodu (obrazu).

4. FORMULACE KRITÉRIA

Zmíněná prostorová představa nám umožňuje formulovat kritérium separability, které, jak bude ukázáno, dovoluje stanovit celou množinu řešení.

Nechť nadrovina ϱ je dána rovnicí

$$x_1 w_1 + x_2 w_2 + \ldots + x_n w_n + w_0 = 0\,; \tag{8}$$

vzdálenost bodu ${}_0(x_i)$ od nadroviny ϱ je dána výrazem

$$\vartheta_0' = \left| \frac{{}_0x_1 w_1 + \ldots + {}_0x_n w_n + w_0}{\sqrt{(w_1^2 + w_2^2 + \ldots + w_n^2 + w_0^2)}} \right|. \tag{9}$$

Pro uvedený problém se jeví výhodné uvažovat vzdálenost včetně její orientace tj.

$$\vartheta_0 = \frac{{}_0x_1 w_1 + \ldots + {}_0x_n w_n + w_0}{\sqrt{(w_1^2 + \ldots + w_n^2 + w_0^2)}}\,. \tag{10}$$

Zavedeme dále

$$\bar{\vartheta}' = \frac{d}{\sqrt{(w_1^2 + \ldots + w_n^2 + w_0^2)}}\,. \tag{11}$$

Je dáno s různých bodů ${}_k(x_i)$, $k = 1, \ldots, s$, o souřadnicích ${}_kx_i$, $i = 1, \ldots, n$; $k = 1, \ldots, s$. Jde o tzv. trénovací množinu a příslušnost těchto bodů do dvou možných skupin je rovněž známa. (Vždy je možné provést přečíslování bodů tak, aby body ${}_k(x_i)$, $k = 1, \ldots, r$, patřily do prvé a body ${}_k(x_i)$, $k = r + 1, \ldots, s$, do druhé skupiny.)

To znamená, že u bodů prvé skupiny požadujeme, aby jejich vzdálenost od nadroviny (8) byla

$$\vartheta_k > \bar{\vartheta}'\,, \quad k = 1, \ldots, r\,, \tag{12}$$

a u druhé skupiny

$$-\vartheta_k > \bar{\vartheta}'\,, \quad k = r + 1, \ldots, s\,. \tag{13}$$

Uvážíme-li vztahy (10) a (11), pak nerovnosti (12) a (13) můžeme přepsat ve tvaru

$${}_kx_1 w_1 + \ldots + {}_kx_n w_n + w_0 > d\,, \tag{14}$$

pro

$$k = 1, \ldots, r$$

a

$$-{}_kx_1 w_1 - \ldots - x_n w_n - w_0 > d \tag{15}$$

pro

$$k = r + 1, \ldots, s\,.$$

Přihlédneme-li k tomu, že všechna ${}_kx_i$, $k = 1, \ldots, s$, $i = 1, \ldots, n$, jsou dána a přepíšeme-li nerovnosti (14) a (15) vhodným způsobem do tvaru (16), pak vidíme, že jde o soustavu lineárních nerovností vzhledem k w_i, $i = 0, 1, \ldots, n$.

$$
(16) \qquad
\begin{aligned}
{}_1x_1w_1 + \ldots + {}_1x_nw_n + w_0 &> d\,, \\
&\vdots \\
{}_rx_1w_1 + \ldots + {}_rx_nw_n + w_0 &> d\,, \\
-{}_{r+1}x_1w_1 - \ldots - {}_{r+1}x_nw_n - w_0 &> d\,, \\
&\vdots \\
-{}_sx_1w_1 - \ldots - {}_sx_nw_n - w_0 &> d\,.
\end{aligned}
$$

Doplňme soustavu nerovností (16) na soustavu rovnic (17) a to tak, že po řadě do levé strany každé nerovnosti soustavy přičteme novou proměnnou y_k, $k = 1, \ldots, s$:

$$
(17) \qquad
\begin{aligned}
{}_1x_1w_1 + \ldots + {}_1x_nw_n + w_0 + y_1 &= d\,, \\
&\vdots \\
{}_rx_1w_1 + \ldots + {}_rx_nw_n + w_0 + y_r &= d\,, \\
-{}_{r+1}x_1w_1 - \ldots - {}_{r+1}x_nw_n - w_0 + y_{r+1} &= d\,, \\
&\vdots \\
-{}_sx_1w_1 - \ldots - {}_sx_nw_n - w_0 + y_s &= d\,.
\end{aligned}
$$

Rozšířená matice soustavy (17) má pak tvar

$$
(18) \qquad
\left[
\begin{array}{cccc|ccccccc|c}
{}_1x_1 & {}_1x_2 \ldots & {}_1x_n & 1 & 1\ 0 & & \ldots & & & 0\ 0 & d \\
\vdots & & & & & & & & & & \\
{}_rx_1 & {}_rx_2 \ldots & {}_rx_n & 1 & 0 \ldots 0\ 1\ 0 & & \ldots & & & 0 & d \\
-{}_{r+1}x_1 & -{}_{r+1}x_2 \ldots & -{}_{r+1}x_n & -1 & 0 \ldots 0\ 0\ 1\ 0 & & \ldots & & & 0 & d \\
\vdots & & & & & & & & & & \\
-{}_sx_1 & -{}_sx_2 \ldots & -{}_sx_n & -1 & 0\ 0 & & \ldots & & & 0\ 1 & d
\end{array}
\right].
$$

Z tvaru matice (18) je okamžitě vidět, že její hodnost je s a tedy soustava (17) má vždy řešení vzhledem k proměnným y_k, $k = 1, \ldots, s$ a s proměnnými w_i, $i = 0, 1, \ldots$ $\ldots, n$ jako parametry. Vzhledem k tomu, jak jsme zavedli proměnné y_k lze tvrdit, že soustava nerovností (16) bude mít alespoň jedno řešení vzhledem k w_i, $i = 0, 1, \ldots, n$, právě tehdy, existuje-li alespoň jeden soubor takových

$$
(19) \qquad y_k < 0\,, \quad k = 1, \ldots, s\,,
$$

který vyhovuje soustavě rovnic (17).

Předpokládejme tedy volbu všech y_k podle (19) a zkoumejme, za jakých okolností je nyní řešitelná soustava (17), ale vzhledem k proměnným w_i, $i = 0, 1, \ldots, n$, a s proměnnými y_k, $k = 1, \ldots, s$, jako (předem podle (19) zvolenými) parametry. Za tím účelem budeme upravovat matici (18) tak, aby její podmatice sestavená z koeficientů soustavy (16) nabyla tvaru diagonální matice. Úprava se provádí v zásadě Gaussovou eliminační metodou.

Výsledný tvar matice získané uvedenou úpravou matice (18) se bude lišit případ od případu v závislosti na tom, jaký je poměr hodnosti nerozšířené matice soustavy (16) (označme p) a hodností matice (18). Jak již bylo konstatováno hodnost matice (18) je vždy s.

Posuďme nejprve případ, že

$$p = s . \tag{a}$$

V tomto případě je zřejmé, že řešení soustavy (17) vzhledem k proměnným w_i, $i = 0, 1, \ldots, n$, je jediné, ovšem závislé na parametrech y_k, $k = 1, \ldots, s$. Je tedy možno parametry y_k zvolit tak, aby splňovaly podmínku (19). Tím bude zaručena existence řešení soustavy (16). Původní problém je pak lineárně separabilní a všechna řešení (tj. nadroviny) dostaneme různými volbami parametrů y_k a jejich dosazením do rovnic pro w_i, $i = 0, 1, \ldots, n$.

V případě

$$p < s \tag{b}$$

lze zřejmě v matici soustavy (16) nalézt právě p lineárně nezávislých řádků. Lze tedy právě p z proměnných w_i, $i = 0, 1, \ldots, n$, vypočítat v závislosti na y_k, $k = 1, \ldots$ $\ldots, s$.

Z řádků, které byly v matici soustavy (16) lineárně závislé (je jich právě $(s - p)$), vznikne tedy $(s - p)$ rovnic v proměnných y_k, $k = 1, \ldots, s$. (Poznamenejme ještě, že těchto rovnic je určitě méně než s, protože $1 \leqq p < s$.)

Položme

$$s - p = q ; \tag{20}$$

pak nově vzniklá soustava rovnic má tvar

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 + \ldots + a_{1s}y_s &= A_1 , \\ &\vdots \\ a_{q1}y_1 + a_{q2}y_2 + \ldots + a_{qs}y_s &= A_q , \end{aligned} \tag{21}$$

kde a_{ij}, $i = 1, \ldots, q$, $j = 1, \ldots, s$, jsou koeficienty vzniklé při aplikaci eliminačního postupu na soustavu (17) a A_i, $i = 1, \ldots, q$, jsou absolutní členy vzniklé při témže postupu. Protože $s > p$ a hodnost matice soustavy (21) je $q = s - p$, má soustava (21) vždy řešení. Aplikujeme tedy na soustavu již uvedený eliminační postup. Jako výsledek dostaneme následující soustavu rovnic (předpokládáme, že vyjádříme právě y_1 až y_q; vhodným přerovnáním a přečíslováním proměnných y_k toho lze dosáhnout)

$$\begin{aligned} y_1 &= b_{1,q+1}y_{q+1} + \ldots + b_{1,s}y_s + B_1 , \\ &\vdots \\ y_q &= b_{q,q+1}y_{q+1} + \ldots + b_{q,s}y_s + B_q . \end{aligned} \tag{22}$$

Vzhledem k tomu, že pro y_k, $k = 1, \ldots, q$, musí platit podmínka (19), musí pro pravé strany rovnic soustavy (22) platit

$$(23) \qquad \begin{aligned} & b_{1,q+1}y_{q+1} + \ldots + b_{1,s}y_s + B_1 < 0\,, \\ & \vdots \\ & b_{q,q+1}y_{q+1} + \ldots + b_{q,s}y_s + B_q < 0\,. \end{aligned}$$

Na soustavu (23) aplikujeme opět eliminační postup a úvahu o platnosti podmínky (19) pro všechna y_k, $k = 1, \ldots, s$. Tímto postupem cyklicky opakovaným dospějeme posléze k jednoduchým podmínkám pro některá y_k. Podmínky budou nerovnosti typu

$$(24) \qquad y_j < G_j\,,$$

$$(25) \qquad y_k > G_k\,,$$

kde G_j, G_k jsou absolutní členy. Pokud jsou všechny podmínky (24) resp. (25) slučitelné s podmínkami (19), lze říci, že soustava (17) je řešitelná s podmínkou (19) a tedy, že je řešitelná i soustava (16) a konečně, že původní problém je lineárně separabilní.

Poznámka 1. Lze snadno nahlédnout, že někdy je možno o řešitelnosti problému rozhodnout, aniž bychom dováděli výpočet do konce. Např. jestliže v některé rovnici soustavy (21) je $A_i \geqq 0$ a všechna $a_{ij} > 0$, $j = 1, \ldots, s$, pak tato rovnice je zřejmě neslučitelná s podmínkou (19) a problém není lineárně separabilní.

Poznámka 2. Podaří-li se problém prokázat jako lineárně separabilní, lze již pouhou volbou různých vhodných souborů y_k, $k = 1, \ldots, s$, a dosazením do příslušných rovnic získat koeficienty w_i, $i = 0, 1, \ldots, n$, nadrovin, které řeší danou úlohu. Čili testem lineární separability je nahrazen proces učení, tj. hledání vhodné nadroviny.

Poznámka 3. Z podmínek (24) resp. (25) lze říci, zda neseparabilita nebyla způsobena nevhodnou volbou pásma necitlivosti d, jinými slovy, pro jaká d by problém byl lineárně separabilní.

5. ZÁVĚR

O odvozeném kritériu lze říci, že dovoluje exaktně stanovit existenci řešení resp. množiny řešení daného problému učení. Jestliže řešení existuje, pak umožňuje jeho nalezení (resp. výběr jednoho z množiny řešení). Vzhledem k charakteru kritéria je patrné, že ho bude lze použít zejména při modelování uvedených procesů na číslicovém počítači.

(Došlo dne 23. ledna 1968.)

[1] Widrow B., Pierce W. H., Angell J. B.: Birth, life and death in microelectronic systems. IRE Trans. *MIL-5* (1961 July), 191–201.
[2] Beneš J.: Kybernetické systémy s automatickou organizací. Academia, Praha 1966.

SUMMARY

The Test of Linear Separability

MILOŠ THOMA

The paper deals with the testing of linear separability. We meet this problem in the application of the treshold logical adaptive circuit in pattern recognition and learning.

The treshold logical adaptive circuit realizes a linear discriminant function. The determination of the form of this function is usually carried out by certain learning methods converging if their solution exists. The paper solves the question of the existence of solution i.e. the existence of the linear discriminant function solving the problem. The criterion that is the result of the paper answers not only the question of the existence of the solution, but it also enables to find out its form.

Ing. Miloš Thoma, Ústav teorie informace a automatizace ČSAV, Vyšehradská 49, Praha 2.