УДК 681.14

В. М. Коцовський (Ужгородський нац. ун-т)

ВЛАСТИВОСТІ ДИХОТОМІЙ, ПОРОДЖЕНИХ ДВОПОРОГОВИМИ НЕЙРОННИМИ ЕЛЕМЕНТАМИ

This paper is concerned with the properties of bithreshold elements. We shall first investigate the sufficient condition of bithreshold separability in n-dimensional space and correct a previously published results. Finally, we shall estimate lower and improve upper bounds for the capacity of bithreshold elements, using reduction to the case of linear separability.

В роботі розглядаються питання пов'язані з двопороговими нейронними елементами. Встановлюються нові та уточнюються відомі достані умови двопорогової віддільності в n-вимірному просторі. Наводяться покращення відомих оцінок двопорогово сепарабельних дихотомій скінченної множини.

Нейронні елементи та нейромережі, побудовані з них знаходять застосування при розв'язуванні широкого кола практичних задач [1]. У зв'язку з обмеженими можливостями класичних порогових елементів інтенсивно вивчалися їх різноманітні узагальнення [2]. Двопорогові нейронні елементи (ДНЕ) з дійсними вагами та порогами (bithreshold neuron) були введені у розгляд багатьма авторами (наприклад у роботах [3-8]). У цих роботах були встановлені оцінки кількості двопорогово сепарабельних дихотомій скінченних множин у просторі $\mathbb{R}^n$. Подальші властивості ДНЕ розглядалися у [9], де були встановлені достатні умови, які забезпечують двопорогову сепарабельність підмножини $\mathbb{R}^n$. У роботах [10-11] з використанням властивостей матриць толерантності та характеристичних векторів досліджувалися питання, пов'язані з можливістю відокремлення вершин n-вимірного гіперкуба за допомогою ДНЕ. У даній роботі наводятьяя контрприклади до основного твердження роботи [9] про двопорогову сепарабельність, встановлюються коректні достатні умови двопорогової сепарабельності, покращуються оцінки кількості двопорогово сепарабельних дихотомій скінченних множин та встановлюється асипмтотична поведеінка цих оцінок. Також показано, що розмірність Вапніка-Червоненкіса [1] (VCDim) класу $\mathcal{F}$ ДНЕ з дійсними вагами задовольняє подвійну нерівність $2n \le \mathrm{VCDim}\,\mathcal{F} \le 5n$.

1. Основні поняття та означення. Нехай $\mathbb{R}^n$ — n-вимірний дійсний евклідів простір. Якщо $\mathbf{w} = (w_1, , \dots, w_n) \in \mathbb{R}^n$, $\mathbf{x} \in (x_1, \dots, x_n) \in \mathbb{R}^n$, то величину скалярного добутку

$$(\mathbf{w}, \mathbf{x}) = \sum_{i=1}^{n} w_i x_i$$

будемо називати зваженою сумою, що відповідає вектору $\mathbf{x}$. При цьому вектор $\mathbf{w}$ будемо називати ваговим вектором. Двопороговим дійсним нейронним елементом з ваговим вектором $\mathbf{w} \in \mathbb{R}^n$, порогами t_1, $t_2 \in \mathbb{R}$ $(t_1 < t_2)$ будемо називати функціональний елемент з n дійсними входами $x_1, x_2, \dots, x_n$ та одним виходом y, поведінка якого описується співвідношеннями:

$$y = \begin{cases} a, & \text{якщо} \ \ t_1 < (\mathbf{w}, \mathbf{x}) < t_2, \\ b, & \text{якщо} \ \ (\mathbf{w}, \mathbf{x}) \le t_1 \text{ або } (\mathbf{w}, \mathbf{x}) \ge t_2. \end{cases}$$

У більшості практично важливих випадків досить покласти $\{a, b\} = \mathbb{Z}_2$ або $\{a, b\} = G_2$, де $\mathbb{Z}_2 = \{0, 1\}$, $G_2 = \{-1, 1\}$. Надалі будуть розглядатися тільки ті двопорогові дійсні нейронні елементи, для яких $\{a, b\} = G_2$. Для цих елементів буде використовуватися скорочене позначення ДНЕ. В термінах [12] зважена сума є вхідним оператором ДНЕ, параметрична функція активації має вигляд

$$f_{t_1,t_2}(x) = \begin{cases} -1, & \text{якщо } t_1 < x < t_2, \\ \ \ 1, & \text{в інших випадках.} \end{cases}$$

Вихідним оператором ДНЕ є тотожній оператор. ДНЕ повністю визначається впорядкованою трійкою $(\mathbf{w}, t_1, t_2)$, яку будемо надалі називати вектором структури або просто структурою ДНЕ. ДНЕ із структурою $(\mathbf{w}, t_1, t_2)$ здійснює розбиття простору $\mathbb{R}^n$ на дві підмножини наступним чином:

$$\mathbb{R}^{n+} = \{x \in \mathbb{R}^n |\, (\mathbf{w}, \mathbf{x}) \le t_1\} \cup \{x \in \mathbb{R}^n |\, (\mathbf{w}, \mathbf{x}) \ge t_2\}, \quad \mathbb{R}^{n-} = \mathbb{R} \setminus \mathbb{R}^{n+}.$$

Дві множини $A^+ \subset \mathbb{R}^n$ і $A^- \subset \mathbb{R}^n$ назвемо двопорогово сепарабельними (д-розділимими), якщо знайдеться такий ДНЕ із вектором структури $(\mathbf{w}, t_1, t_2)$, що $A^+ \subset \mathbb{R}^{n+}$ і $A^- \subset \mathbb{R}^{n-}$. У цьому випадку будемо казати, що ДНЕ із структурою $(\mathbf{w}, t_1, t_2)$ здійснює двопорогову дихотомію (д-розбиття) (A^+, A^-) множини A на дві множини A^+ і A^-, що не перетинаються. Якщо ДНЕ має структуру $(\mathbf{w}, t_1, t_2)$, то для довільної множини $A \subset \mathbb{R}^n$ її підмножини $A^+ = A \cap \mathbb{R}^{n+}$, $A^- = A \cap \mathbb{R}^{n-}$ є д-розділимими. Легко бачити, що д-розділимість множин є узагальненням лінійної розділимості. Вже у одновимірному просторі можна навести приклад д-розділимих множини, яка не є лінійно розділимими (наприклад $A^- = \{0\}$, $A^+ = \{-1, 1\}$).

У множині ДНЕ, які здійснюють д-розбиття множини значний інтерес становлять ті ДНЕ, вагові вектори яких породжують лінійний порядок на множині A. Має місце наступна лема.

Лема 1. *Якщо ДНЕ із структурою $(\mathbf{w}, t_1, t_2)$ здійснює д-розбиття не більше, ніж зліченної, обмеженої множини $A \subset \mathbb{R}^n$ і справджуються нерівності*

$$t_1 < \inf_{\mathbf{x} \in A^-} (\mathbf{w}, \mathbf{x}), \quad \sup_{\mathbf{x} \in A^-} (\mathbf{w}, \mathbf{x}) < t_2 \tag{1}$$

то знайдеться ДНЕ із структурою $(\mathbf{w}', t'_1, t'_2)$, який здійснює таке саме д-розбиття множини A, причому $\forall \mathbf{x}, \mathbf{y} \in A \quad (\mathbf{w}', \mathbf{x}) = (\mathbf{w}', \mathbf{y}) \Rightarrow \mathbf{x} = \mathbf{y}$ і значення жодної зваженої суми не співпадає із порогами t'_1 або t'_2.

Доведення. Нехай

$$\delta = \frac{1}{2} \min \left\{ \inf_{x \in A^-} (\mathbf{w}, \mathbf{x}) - t_1,\, t_2 - \sup_{x \in A^-} (\mathbf{w}, \mathbf{x}), 1, \|\mathbf{w}\| \right\}, \quad d = \max \left\{ \sup_{\mathbf{x} \in A} \|\mathbf{x}\|, 1 \right\}.$$

Пронумеруємо за допомогою натуральних чисел усі невпорядковані пари $\{\mathbf{x}^k, \mathbf{y}^k\}$ векторів, кожний з яких належать множині A. Будемо індуктивно будувати послідовність вагових векторів $\{\mathbf{w}^k\}$ наступним чином. Покладемо $\mathbf{w}^1 = \mathbf{w}$. Будемо вважати, що вектори $\mathbf{w}^1, \ldots, \mathbf{w}^k$ вже побудовані і $c_k = \min\{ (\mathbf{w}^{j+1}, \mathbf{x}^j - \mathbf{y}^j) |\ 1 \le j < k\}$, $c'_k = \min\left\{ \left(\mathbf{w}^k, \mathbf{x}^j - \mathbf{y}^j\right) \middle|\, 1 \le j < k \right\}$. Якщо $\left(\mathbf{w}^k, \mathbf{x}^k\right) \ne \left(\mathbf{w}^k, \mathbf{y}^k\right)$, то покладемо $\mathbf{w}^{k+1} = \mathbf{w}^k$. У протилежному випадку $\mathbf{w}^{k+1} = \mathbf{w}^k + \alpha_k \tilde{\mathbf{x}}^k$, де $\left\|\tilde{\mathbf{x}}^k\right\| = 1$,

$\left(\mathbf{x}^k, \tilde{\mathbf{x}}^k\right) = 0$, $\left(\mathbf{y}^k, \tilde{\mathbf{x}}^k\right) \neq 0$ (ми вважаємо, що $\mathbf{y}^k \neq \mathbf{0}$, інакше можна поміняти місцями елементи пари $\left\{\mathbf{x}^k, \mathbf{y}^k\right\}$) і

$$\alpha_k = \delta \cdot \left(2^{k+1} \cdot d \cdot \max_{1 \le j < k} \left\|\mathbf{x}^j - \mathbf{y}^j\right\|\right)^{-1} \cdot \min\left\{c_k, c'_k\right\} \cdot \min\left\{1, \left\|\mathbf{w}^k\right\|^{-1}\right\}.$$

З правила побудови вектора $\mathbf{w}^{k+1}$ випливає, що

$$\left(\mathbf{w}^{k+1}, \mathbf{x}^k\right) \neq \left(\mathbf{w}^{k+1}, \mathbf{y}^k\right),$$

$$\left|\left(\mathbf{w}^{k+1}, \mathbf{x}^j\right) - \left(\mathbf{w}^{k+1}, \mathbf{y}^j\right)\right| \ge \frac{1}{2}\left|\left(\mathbf{w}^{j+1}, \mathbf{x}^j\right) - \left(\mathbf{w}^{j+1}, \mathbf{y}^j\right)\right|, \quad 1 \le j < k$$

і $\forall \mathbf{x} \in A \quad \mathbf{x} \in A^- \Leftrightarrow t_1 + \dfrac{\delta}{2} < \left(\mathbf{w}^{k+1}, \mathbf{x}\right) < t_2 - \dfrac{\delta}{2}$. Оскільки $\left|\left\|\mathbf{w}^{k+1}\right\| - \left\|\mathbf{w}^k\right\|\right| < \delta \cdot 2^{-k-1}$ і $\left\|\mathbf{w}^{k+1}\right\| > \left\|\mathbf{w}\right\|/2$, то існує $\lim\limits_{k\to\infty} \mathbf{w}^k = \mathbf{w}'$, причому $\mathbf{w}' \neq \mathbf{0}$ (у випадку скінченності множини A починаючи з індексу $k = \mathrm{Card}\, A\, (\mathrm{Card}\, A - 1)/2$ вектори $\mathbf{w}^k$ не змінюються). ДНЕ із вектором структури $(\mathbf{w}', t_1 + \delta, t_2 - \delta)$, очевидно, задовольняє умови леми.

Зауваження 1. Умова обмеженості та зліченності множини A та умова (1) є суттєвими. Можна навести приклади д-розділимих множин, які не задовольняють рівно одну із вказаних умов, для яких лема 1 не справджується.

Зауваження 2. Легко показати, що умова (1) разом з обмеженістю множини A є достатньою для існування цілочислового вектора структури ДНЕ, який здійснює д-розбиття множини A (спочатку треба побудувати відповідну структуру з раціональними коефіцієнтами, а потім помножити її на спільне кратне знаменників усіх коефіцієнтів). Проте можна навести приклади зліченної множини, на якій жодний цілочисловий ваговий вектор вже не задає лінійний порядок.

2. Умови д-розділимості. Для класичного перцептрона було встановлено зв'язок між лінійною розді-лимістю множин у просторі R^n та властивостями опуклих оболонок цих множин: скінченні множини точок A і B лінійно розділимі тоді і тільки тоді, коли $\mathrm{conv}\, A \cap \mathrm{conv}\, B = \varnothing$. У випадку ДНЕ зв'язок між д-розділимістю множин і їх лінійними оболонками значно складніший.

Лема 2. *Якщо ДНЕ із структурою* $(\mathbf{w}, t_1, t_2)$ *здійснює д-розбиття множини на множини* A^+ *і* A^-, *то*

$$A^+ \cap \mathrm{conv}\, A^- = \varnothing \tag{2}$$

і множини A^+ *і* $\mathrm{conv}\, A^-$ *є д-розділимими.*

Доведення. Нехай $\mathbf{y}$ — довільний елемент множини A. Тоді за означенням опуклої оболонки

$$\mathbf{y} = \sum_{i=1}^{m} \lambda_i \mathbf{x}^i, \ \lambda_i \in [0,1], \ \sum_{i=1}^{m} \lambda_i = 1, \ \mathbf{x}^i \in A^-, \ i = 1, \ldots, m.$$

Тому

$$t_1 = \sum_{i=1}^{m} \lambda_i t_1 < \sum_{i=1}^{m} \lambda_i \left(\mathbf{w}, \mathbf{x}^i\right) = (\mathbf{w}, \mathbf{y}) < \sum_{i=1}^{m} \lambda_i t_2 = t_2.$$

З останньої рівності випливає, що $\mathbf{y} \notin A^+$ і ДНЕ з структурою $(\mathbf{w}, t_1, t_2)$ здійснює д-розбиття множини $A^+ \cup \operatorname{conv} A^-$ на A^+ і $\operatorname{conv} A^-$. Твердження обернене до твердження леми 2 не є вірним у випадку $n \geq 2$.

У зв'язку з вищесказаним значний інтерес становить встановлення додаткових умов, при виконанні яких з (2) вже буде випливати д-розділимість множин A^+ і A^-. Перші кроки у цьому напрямку були зроблені у [9]. На жаль, доведення теореми 1 у [9] містило логічні помилки і як далі буде показано, основне твердження теореми невірне.

Вектор $\mathbf{y} \in \mathbb{R}^n$ називається афінною комбінацією n-вимірних дійсних векторів $\mathbf{x}^1, \ldots, \mathbf{x}^k$, якщо

$$\mathbf{y} = \alpha_1 \mathbf{x}^1 + \ldots + \alpha_k \mathbf{x}^k, \quad \sum_{i=1}^{k} \alpha_i = 1, \quad \alpha_i \in \mathbb{R}, \quad i = 1, \ldots, k. \tag{3}$$

Множину всіх точок, кожна з яких є афінною комбінацією деяких точок множини позначимо через $\operatorname{Aff}(X)$.

Теорема 1. *Нехай A^+ — не більш, ніж зліченний компакт у просторі $\mathbb{R}^n$ і $A^- = \left\{\mathbf{x}^1, \ldots, \mathbf{x}^k\right\}$ — множина лінійно незалежних векторів, $k \leq n$. Тоді якщо знайдеться такий вектор $\mathbf{x}^l \in A^-$, що для довільного $\mathbf{y} \in A^+ \cap \operatorname{Aff}(A^-)$ коефіцієнт α_l у розкладі (3) задовольняє умову*

$$\alpha_l \in (-\infty, 0) \cup (1, +\infty), \tag{4}$$

то множини A^+ і A^- є д-розділимими.

Доведення. Покажемо спочатку, що можна провести гіперплощину $H_{\mathbf{w}} = \left\{\mathbf{x} \in \mathbb{R}^n | \ (\mathbf{w}, \mathbf{x}) = c\right\}$, $\mathbf{w} \neq \mathbf{0}$, $c \neq 0$, якій належать усі точки множини A^- і для всіх $\mathbf{y} \in A^+$ з того, що $\mathbf{y} \in H_{\mathbf{w}} \cap A^+$ випливає, що $\mathbf{w}$ потрапляє у лінійну оболонку множини A^-. Для $k = n$ існування такої гіперплощини очевидне. Нехай $k < n$, $A^+ \setminus \langle A^- \rangle = \{\mathbf{y}^1, \mathbf{y}^2, \ldots\}$, де $\langle A^- \rangle$ — лінійна оболонка множини A^-, $D = \max\left\{ \|\mathbf{y}\| \mid \mathbf{y} \in A^+\right\} \neq 0$ (вважаємо, що A^+ містить відмінні від нуля елементи, інакше д-розділимість зводиться до лінійної розділимості). Побудуємо збіжну послідовність векторів $\{\mathbf{w}^m\}$, які задовольняють наступні умови:

$$\forall \mathbf{x} \in A^- \ \ (\mathbf{w}^m, \mathbf{x}) = 1, \quad \left(\mathbf{w}^m, \mathbf{y}^j\right) = c_{mj}, \quad |c_{mj} - 1| \geq d_j, \ j = 1, \ldots, m-1. \tag{5}$$

Нехай $\mathbf{w}^1$ — вектор нормалі довільної гіперплощини $H_1 = \{\mathbf{x} \in \mathbb{R}^n | \ (\mathbf{w}^1, \mathbf{x}) = 1\}$ такої, що $A^- \subset H_1$. Існування H_1 випливає із лінійної незалежності елементів множини A^-. Будемо вважати, що вектор $\mathbf{w}^m$ і числа d_j, $1 \leq j < m$ вже визначені і покажемо, як можна на їх основі побудувати $\mathbf{w}^{m+1}$ і d_m. Якщо $(\mathbf{w}^m, \mathbf{y}^m) \neq 1$, то покладемо $\mathbf{w}^{m+1} = \mathbf{w}^m$. Якщо $(\mathbf{w}^m, \mathbf{y}^m) = 1$, то знайдемо такий вектор $\mathbf{v}^m$, що $(\mathbf{v}^m, \mathbf{x}^j) = 0, \quad 1 \leq j \leq k$, $(\mathbf{v}^m, \mathbf{y}^m) = 1$. Покладемо $\mathbf{w}^{m+1} = \mathbf{w}^m + \beta_m \mathbf{v}^m$, де $\beta_m = \left(2^m \cdot \max\left\{\|\mathbf{v}^m\|, 1\right\} \cdot D\right)^{-1} \cdot \min\limits_{1 \leq j < m} d_j$ (якщо $m = 1$, то вважаємо, що $\min\limits_{1 \leq j < m} d_j = 1$), $d_m = \frac{1}{2}\left|(\mathbf{w}^m, \mathbf{y}^m) - 1\right|$. Тоді для вектора $\mathbf{w}^{m+1}$ виконуються умови (5). Оскільки $\|\mathbf{w}^{m+p} - \mathbf{w}^m\| < 2^{-m+1} d_1 D$, то послідовність $\{\mathbf{w}^m\}$ збігається до вектора $\mathbf{w} \in R^n$, для якого $\forall \mathbf{x} \in A^- \ \ (\mathbf{w}^m, \mathbf{x}) = 1, \quad \forall j \in N \quad |(\mathbf{w}^m, \mathbf{y}^j) - 1| \geq d_j$ (для випадку, коли $\operatorname{Card}\left(A^+ \setminus \langle A^- \rangle\right) = m < \aleph_0$ досить покласти $\mathbf{w} = \mathbf{w}^{m+1}$).

Отже, гіперплощина $H_{\mathbf{w}}$ задовольняє поставленим вимогам. Тоді для множин $H_{\mathbf{w}}$ і A^+ можливими є два випадки.

Випадок 1. $A^+ \cap H_{\mathbf{w}} = \varnothing$. За теоремою Вейєрштрасса неперервна функція $f(\mathbf{x}) = \frac{1}{2}\left|(\mathbf{w}, \mathbf{x}) - 1\right|$ досягає на A^+ своє найменше відмінне від нуля значення δ. Тоді ДНЕ із структурою $(\mathbf{w}, 1-\delta, 1+\delta)$ здійснює д-розбиття множин A^+ і A^-.

Випадок 2. $A_H^+ = A^+ \cap H_{\mathbf{w}} \neq \varnothing$. Покажемо, що у цьому випадку множині A_H^+ належать тільки ті точки множини A^+, які потрапляють у $\mathrm{Aff}\,(A^-)$. Нехай $\mathbf{y} \in A_H^+$. З правила вибору гіперплощини $H_{\mathbf{w}}$ випливає, що $\mathbf{y} = \alpha_1\mathbf{x}^1 + \ldots + \alpha_k\mathbf{x}^k$. Тоді $\mathbf{w}$ є афінною оболонкою точок множини A^-, оскільки

$$\sum_{i=1}^{k} \alpha_i = \sum_{i=1}^{k} \alpha_i \left(\mathbf{w}, \mathbf{x}^i\right) = (\mathbf{w}, \mathbf{y}) = 1.$$

Крім того, легко бачити, що принаймні один коефіцієнт α_i у розкладі (3) є від'ємним, оскільки $\mathbf{y} \notin \mathrm{conv}\, A^-$. Доповнимо множину A^- до базису простору $\mathbb{R}^n$. Нехай $\left\{\mathbf{x}^1, \ldots, \mathbf{x}^k, \mathbf{x}^{k+1}, \ldots, \mathbf{x}^n\right\}$ — отриманий базис. Тоді кожний елемент у простору $\mathbb{R}^n$ можна записати у такому вигляді $\mathbf{y} = \alpha_1\mathbf{x}^1 + \ldots + \alpha_l\mathbf{x}^l + \ldots + \alpha_n\mathbf{x}^n$. Із обмеженості множини A^+ випливає існування такого $\alpha_{\max}$, що для всіх $\mathbf{y} \in A^+$ $|\alpha_j| < \alpha_{\max}$, $j = 1, \ldots, n$. Нехай $A_l^+ = \{\mathbf{y} \in A^+ |\ \alpha_i \in [0,1]\}$. Легко бачити, що множина A_l^+ замкнута і $A_l^+ \cap H_{\mathbf{w}} = \varnothing$. Для кожного $\mathbf{z} \in A_l^+$ $(\mathbf{w}, \mathbf{z}) \neq 1$. Тому знайдеться така відкрита куля $B(\mathbf{z}, r_{\mathbf{z}})$ з центром у точці $\mathbf{z}$ і радіусом $r_{\mathbf{z}}$, що її замикання $B[\mathbf{z}, r]$ немає з гіперплощиною $H_{\mathbf{w}}$ спільних точок. За лемою Бореля відкрите покриття замкнутої множини A_l^+ містить скінченне підпокриття $\{B(\mathbf{z}^1, r_1), \ldots, B(\mathbf{z}^s, r_s)\}$ і величина $\delta = \frac{1}{2}\min\limits_{1 \le j \le s} \min\{|(\mathbf{w}, \mathbf{z}) - 1| \,|\, \mathbf{z} \in B[\mathbf{z}^s, r_s]\}$ є строго додатною. Покажемо, що знайдуться таке $\varepsilon > 0$, що

$$\text{для всіх } \mathbf{y} \in A^+ \quad \left|\alpha_l - \frac{1}{2}\right| \ge \frac{1}{2} + \varepsilon \text{ або } \left|(\mathbf{w}, \mathbf{y}) - 1\right| \ge 2\delta. \tag{6}$$

Припустимо протилежне. Тоді можна побудувати послідовність $\{\mathbf{y}^m\}$, яка задовольняє умову

$$\text{для всіх } m \in \mathbb{N} \quad \left|\alpha_{ml} - \frac{1}{2}\right| < \frac{1}{m} \quad \text{і} \quad \left|(\mathbf{w}, \mathbf{y}^m) - 1\right| < \frac{1}{m}.$$

З виконання першої нерівності випливає, що послідовність $\{\mathbf{y}^m\}$ містить збіжну послідовність, яка збігається до деякого $\mathbf{y}^* \in A_l^+$ для якого $|(\mathbf{w}, \mathbf{y}^*) - 1| = 0$. Це суперечить тому, для всіх $\mathbf{y} \in A_l^+$ $|(\mathbf{w}, \mathbf{y}) - 1| \ge 2\delta$. Виберемо вектор $\mathbf{u} \in \mathbb{R}^n$ таким чином, щоб $(\mathbf{u}, \mathbf{x}^j) = 0$ при $j \in \{1, \ldots, n\}$, $j \neq l$ і $\left(\mathbf{u}, \mathbf{x}^l\right) = 1$. Розглянемо вектор $\tilde{\mathbf{w}} = \mathbf{w} + \gamma\mathbf{u}$, де

$$\gamma < \frac{\delta}{\max\{\alpha_{\max}, \varepsilon + 1\}}.$$

Для всіх $\mathbf{x} \in A^-$ $(\tilde{\mathbf{w}}, \mathbf{x}) \in \{1, 1+\gamma\}$. Якщо $\mathbf{y} \in A^+$, то згідно до (6) $|(\mathbf{w}, \mathbf{y}) - 1| \ge 2\delta$ або $\alpha_l \in (-\infty, \varepsilon] \cup [1+\varepsilon, +\infty)$. У першому випадку маємо $|(\tilde{\mathbf{w}}, \mathbf{y}) - 1| \ge \delta \ge \gamma(1+\varepsilon)$. У другому випадку бачимо, що $(\tilde{\mathbf{w}}, \mathbf{y}) \le 1 - \gamma\varepsilon$ або $(\tilde{\mathbf{w}}, \mathbf{y}) \ge 1 + \gamma(1+\varepsilon)$. Тому ДНЕ з вектором структури $(\tilde{\mathbf{w}}, 1-\gamma\varepsilon, 1+\gamma(1+\varepsilon))$ д-розділяє множин A^+ і A^-. Теорема доведена.

Наслідок. *В умовах теореми 1 д-розділимим є множини A^+ і* $\operatorname{conv} A^-$.

Зауваження. Лінійна незалежність векторів множини A^- і умова (4) є суттєвими. Покажемо це на наступному прикладі. Нехай

$$A^+ = \{(1,1,1,0,0), (1,-1,-1,0,0), (-1,1,-1,0,0), (-1,-1,1,0,0)\},$$

$$A^- = \{(1,1,-1,0,0), (1,-1,1,0,0), (-1,1,1,0,0), (-1,-1,-1,0,0)\}.$$

Множини A^+ і A^- не є д-розділимими, оскільки добре відомо [11], що не є д-розділимими у $\mathbb{R}^3$ відповідні множини вершин одиничного гіперкуба $\{-1,1\}^3$, координати яких співпадають з першими трьома координатами заданих векторів. Слід зазначити, що наведений приклад є контрприкладом до теореми 1 у [9], у якій стверджувалося, що достатньою умовою д-розділимості скінченних множин A^+ і A^- у просторі $\mathbb{R}^n$ є умови $\operatorname{Card} A^- \leq n-1$ і $A^+ \cap \operatorname{conv} A^- = \varnothing$, обидві з яких виконуються у даному прикладі.

Обмеження на потужність множини A^- $(\operatorname{card} A \leq \aleph_0)$ при доведенні теореми 1 використовувалося лише при побудові гіперплощини $\mathbf{w}$. Можна відмовитися від нього, наклавши на координати векторів "сильніші" обмеження, ніж ті, які даються умовою (4).

Теорема 2. *Нехай A^+ — компакт у просторі $\mathbb{R}^n$ і $A^- = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$ — множина лінійно незалежних векторів, $k \leq n$. Тоді якщо знайдеться такий вектор $\mathbf{x}^l \in A^-$ і таке поповнення $\{\mathbf{x}^1, \ldots, \mathbf{x}^k, \mathbf{x}^{k+1}, \ldots, \mathbf{x}^n\}$ множини A^- до базису простору $\mathbb{R}^n$, що для довільного $\mathbf{y} \in A^+$ його координата α_l у цьому базисі задовольняє умову (4), то множини A^+ і* $\operatorname{conv} A^-$ *є д-розділимими.*

Доведення. Так само як при доведенні теореми 1, використовуючи еквівалентність збіжності векторів $\mathbb{R}^n$ і покоординатної збіжності, можна довести існування такого $\varepsilon > 0$, що

$$\forall \mathbf{y} \in A^+ \;\; \alpha_l \in (-\infty, -2\varepsilon] \cup [1 + 2\varepsilon, +\infty).$$

Із обмеженості множини A^+ випливає існування такого $\alpha_{\max}$, що для всіх $\mathbf{y} \in A^+$ $|\alpha_j| < \alpha_{\max}$, $j = 1, \ldots, n$. Нехай δ додатне число, яке задовольняє нерівність $\delta \leq \dfrac{\varepsilon}{n\alpha_{\max}}$. Тоді завжди можна вибрати вектор $\mathbf{w} \in \mathbb{R}^n$ таким чином, щоб $(\mathbf{w}, \mathbf{x}^l) = 1$, $(\mathbf{w}, \mathbf{x}^j) = \delta$, $j \in \{1, \ldots, l-1, l+1, \ldots, k\}$, $(\mathbf{w}, \mathbf{x}^j) = 0$, $k < j \leq n$. Тоді для всіх $\mathbf{y} \in A^+$ $(\mathbf{w}, \mathbf{y}) \leq -\varepsilon$ або $(\mathbf{w}, \mathbf{y}) \geq 1 + \varepsilon$. Тому ДНЕ з вектором структури $(\mathbf{w}, 0, 1+\varepsilon)$ д-розділяє множин A^+ і A^-. Для остаточного доведення теореми потрібно ще застосувати до A^+ і A^- лему 2.

У теоремі 2 д-розділимість випливала з виконання умови (4) для координат точок множини A^+ у розкладі по елементам деякого базису, який містить у собі множину A^-. Можна показати, що аналогічна теорема має місце і у термінах розкладу по елементах множини A^+.

Теорема 3. *Нехай A^- — компакт у просторі $\mathbb{R}^n$ і $A^+ = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$ — множина лінійно незалежних векторів, $k \leq n$. Тоді якщо знайдеться такий вектор $\mathbf{x}^l \in A^+$ і таке поповнення $\{\mathbf{x}^1, \ldots, \mathbf{x}^k, \mathbf{x}^{k+1}, \ldots, \mathbf{x}^n\}$ множини A^+ до базису простору $\mathbb{R}^n$, що для довільного $\mathbf{y} \in A^-$ його координата α_l у цьому базисі задовольняє умову $\alpha_l \in (0,1)$ то множини A^+ і* $\operatorname{conv} A^-$ *є д-розділимими.*

Доведення. Так само як при доведенні теореми 2, використовуючи еквівалентність збіжності векторів $\mathbb{R}^n$ і покоординатної збіжності, можна довести існування такого $\varepsilon > 0$, що для всіх $\mathbf{y} \in A^-$ $\alpha_l \in [2\varepsilon, 1 - 2\varepsilon]$ і такого $\alpha_{\max}$, що для всіх $\mathbf{y} \in A^-$ $|\alpha_j| < \alpha_{\max}$, $j = 1, \ldots, n$. Нехай δ — додатне число, яке задовольняє нерівність $\delta n \alpha_{\max} \leq \varepsilon$. Тоді можна вибрати вектор $\mathbf{w} \in \mathbb{R}^n$ таким чином, щоб $\left(\mathbf{w}, \mathbf{x}^l\right) = 1$, $\left(\mathbf{w}, \mathbf{x}^j\right) = 0$, $j \in \{1, \ldots, l-1, l+1, \ldots, n\}$. Тоді для всіх $\mathbf{y} \in A^-$ $\varepsilon \leq (\mathbf{w}, \mathbf{y}) \leq 1 - \varepsilon$. Тому ДНЕ з вектором структури $(\mathbf{w}, \varepsilon, 1-\varepsilon)$ д-розділяє множин A^+ і $\operatorname{conv} A^-$.

3. Числові оцінки розпізнавальної здатності ДНЕ. Введення у розгляд ДНЕ та їх дослідження мотивується у літературі більш потужними можливостями цих елементів порівняно із звичайними пороговими елементами по розпізнаванню належності точок у $\mathbb{R}^n$ до одного з двох заданих класів. У роботах [3-5] було встановлено оцінки кількості дихотомій скінченної множини у дійсному векторному просторі, які можна отримати за допомогою ДНЕ. У цьому параграфі буде показано, що використовуючи техніку теорії лінійних нерівностей, можна отримати уточнення цих оцінок.

Нехай $D_1(A)$ — кількість способів розбиття точок скінченної множини А на два класи за допомогою звичайного одношарового перцептрона (порогового елемента). Надалі будемо вважати, що $A \subset \mathbb{R}^n$ і $\operatorname{Card} A = m$. Нехай $D_1(m,n) = \max\{D_1(A) \mid A \subset \mathbb{R}^n,\ \operatorname{Card} A = m\}$. Добре відомо [1], що

$$D_1(m,n) = \begin{cases} 2\sum\limits_{i=0}^{n} C_{m-1}^i, & \text{якщо } m > n+1, \\ 2^m, & \text{якщо } m \leq n+1, \end{cases} \tag{7}$$

де C_{m-1}^i — біномні коефіцієнти, причому $D_1(A) = D_1(m,n)$ тоді і тільки тоді, коли усі точки множини A знаходяться у загальному положенні (через жодні $n+1$ точки не можна провести гіперплощину). Нехай $D_2(A)$ — кількість різних д-розбитів (A^+, A^-) скінченної m-елементної множини $A \subset \mathbb{R}^n$,

$$D_2(m,n) = \max\{D_2(A) \mid\ A \subset \mathbb{R}^n,\ \operatorname{Card} A = m\}.$$

Покажемо, що (7) можна використати для того, щоб отримати оцінки для$D_2(m,n)$, причому верхня оцінка значно покращує оцінку, отриману в роботі [5]. Надалі без додаткових застережень будемо вважати, що $A \subset \mathbb{R}^n$ і $\operatorname{Card} A = m$.

Твердження 1. *При $m > n$ має місце нерівність*

$$D_2(m,n) < 2\sum_{i=0}^{n+1} C_{2m-1}^i. \tag{8}$$

Якщо елементи множини A знаходяться у загальному положенні, то

$$D_2(A) \geq 2\sum_{i=0}^{n+1} C_{m-1}^i - 1. \tag{9}$$

Доведення. Для кожного вектора $\mathbf{x} = (x_1, \ldots, x_n) \in A$ побудуємо два $n+2$-вимірні вектори $\mathbf{x}' = (-x_1, \ldots, -x_n, 1, 0)$ і $\mathbf{x}'' = (x_1, \ldots, x_n, 0, -1)$. Якщо ДНЕ із структурою $(\mathbf{w}, t_1, t_2)$ здійснює д-розбиття (A^+, A^-) множини , то гіперплощина $w_1x_1 + \ldots + w_nx_n + t_1x_{n+1} + t_2x_{n+2} = 0$ здійснює лінійне розбиття (B^+, B^-) множини $B = \{\mathbf{x}' \,|\, \mathbf{x} \in A\} \cup \{\mathbf{x}'' \,|\, \mathbf{x} \in A\}$ і $B_1^- = \{\mathbf{x}' \mid \mathbf{x} \in A^-\} \cup \{\mathbf{x}'' \mid \mathbf{x} \in A^-\} \subset B^-$, $B_1^+ = \{\mathbf{x}' \,|\, \mathbf{x} \in A,\ (\mathbf{w}, \mathbf{x}) \leq t_1\} \cup \{\mathbf{x}'' \,|\, \mathbf{x} \in A,\ (\mathbf{w}, \mathbf{x}) \geq t_2\} \subset B^+$. Тому кількість різних д-розбиттів множини A не перевищує кількість лінійних розбиттів множини B (на-справді має місце строга нерівність, бо B_1^+ - власна підмножина множини B^+). Відомо [2], що кількість лінійних розбиттів множини можна обчислити по формулі (7), зменшивши на 1 кількість доданків. Звідси випливає справедливість нерівності (8).

Оцінимо порядок росту правої частини у (8). Відомо [13], що при $l > n + 1$

$$\sum_{i=0}^{n} C_{l-1}^{i} \leq 1,5\ \frac{l^n}{n\,!}$$

З урахуванням останньої нерівності маємо, що

$$D_2\left(m, n\right) < 3 \cdot \frac{\left(2m\right)^{n+1}}{(n+1)!} = O\left(m^{n+1}\right). \tag{10}$$

Для порівняння можна нагадати [2], що $D_1\left(m, n\right) = O\left(m^n\right)$. Оцінка (10) є значно кращою, ніж верхня оцінка, наведена у роботі [5], де фактично показано, що $D_2\left(m, n\right) < O\left(m^{2n+1}\right)$.

Перейдемо до встановлення нижньої оцінки. Якщо при встановленні верхньої оцінки нами фактично було здійснено перехід від простору $\mathbb{R}^n$ до простору $\mathbb{R}^{n+2}$, то для з'ясування нижньої оцінки перейдемо від $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ до $\tilde{\mathbf{x}} = (x_1, \ldots, x_n, -1) \in \mathbb{R}^{n+1}$. Нехай $\tilde{\mathbf{w}} = (w_1, \ldots, w_n, t_1)$. Тоді

$$\forall \mathbf{x} \in A \ \ \mathbf{x} \in A^- \Leftrightarrow 0 < (\tilde{\mathbf{w}}, \tilde{\mathbf{x}}) < t_2 - t_1.$$

Очевидно, що якщо ДНЕ із структурою $(\mathbf{w}, t_1, t_2)$ здійснює д-розбиття множини A, то гіперплощина $w_1x_1 + \ldots + w_nx_n + t_1x_{n+1} = t_2 - t_1$ розбиває множину $C = \{\tilde{\mathbf{x}} \,|\, \mathbf{x} \in A\}$ на дві множини $C^+ = \{\tilde{\mathbf{x}} \,|\ (\tilde{\mathbf{w}}, \tilde{\mathbf{x}}) \geq t_2 - t_1\}$ і $C^- = C \backslash C^+$. Тоді з урахуванням (7) і того, що елементи множини C вже не знаходяться у загальному положенню отримуємо (9). Твердження доведене.

Розглянемо останній доданок у (9). Легко бачити, що при фіксованому n він є величиною порядку $O\left(m^{n+1}\right)$. Тому нижня оцінка має той самий асимптотичний порядок росту, що й верхня оцінка (10).

Зауваження. Нижня оцінка (9) за допомогою інших методів була встановлена у роботах [3-5] для двопорогових елементів більш загального вигляду. Слід зазначити, що верхню оцінку (8) можна дещо покращити. При цьому будуть встановлені факти, які самі по собі становлять певний інтерес. Уведемо у розгляд поняття порогового 3-розбиття, породженого двома паралельними гіперплощинами. Назвемо впорядковану трійку (A_1, A_2, A_3) пороговим 3-розбиттям множини $A = A_1 \cup A_2 \cup A_3$, породженим вектором $\mathbf{w} \in \mathbb{R}^n$ і порогами $t_1, t_2 \in \mathbb{R}$, якщо

$$A_1 = \{\mathbf{x} \in A |\ (\mathbf{w}, \mathbf{x}) \leq t_1\},\ A_2 = \{\mathbf{x} \in A |\, t_1 < (\mathbf{w}, \mathbf{x}) < t_2\},\ A_3 = A \setminus (A_1 \cup A_2).$$

Нехай $T(A)$ — кількість різних порогових 3-розбиттів множини A і

$$T(m,n) = \max\{T(A)\,|A \subset \mathbb{R}^n,\ \mathrm{Card}\,A = m\}.$$

Лема 3. *Має місце формула*

$$T(m,1) = m^2 + 3m - 1. \tag{11}$$

Доведення. Для доведення (11) досить додати кількість 3-розбиттів, для яких $w_1 = 1$, та кількість 3-розбиттів з $w_1 = -1$ і відняти від суми кількість 3-розбиттів, які входять у обидва доданки. Очевидно, що останніх є рівно 3 ($(A,\varnothing,\varnothing)$, $(\varnothing,A,\varnothing)$ і $(\varnothing,\varnothing,A)$) Елементи множини A розбивають дійсну вісь на $m+1$ проміжок, з яких можна вибирати значення порогів $t_1,\ t_2$, причому t_1 і t_2 можуть потрапляти у один і той самий проміжок. Тому є $\bar{C}_{m+1}^2$ 3-розбиттів, для яких $w_1 = 1$, де $\bar{C}_n^r = C_{n+r-1}^r$ — кількість комбінацій з повтореннями з n об'єктів по r. Така сама кількість 3-розбиттів з $w_1 = -1$, що й доводить формулу (11).

Лема 4. *При $n > 1$ має місце нерівність*

$$T(m+1,n) \le T(m,n) + 2T(m,n-1), \tag{12}$$

при чому рівність виконується тоді і тільки тоді, коли точки знаходяться у загальному положенні.

Доведення. Скористаємося методом, запропонованим у [2] для доведення (7). Нехай $A = \{\mathbf{x}^1,\ldots,\mathbf{x}^m,\mathbf{x}^{m+1}\}$. Позначимо через d кількість порогових 3-розбиттів множини $\tilde{A} = \{\mathbf{x}^1,\ldots,\mathbf{x}^m\}$, для яких одна з гіперплощин проходить через точку $\mathbf{x}^{m+1}$. Для кожного такого 3-розбиття знайдеться не більше двох нових порогових 3-розбиттів, одній з множин яких належить точка $\mathbf{x}^{m+1}$ (якщо точки знаходяться у загальному положенні, то так само, як у [2] можна показати, що таких 3-розбиттів буде рівно 2). Для всіх інших $T(m,n) - d$ порогових 3-розбиттів множини $\tilde{A}$ точка $\mathbf{x}^{m+1}$ потрапляє у одну з трьох множин $\tilde{A}_1$, $\tilde{A}_2$ чи $\tilde{A}_3$. Тому

$$T(m+1,n) \le 2d + T(m,n) - d = T(m,n) + d. \tag{13}$$

Для підрахунку d зауважимо, що обмеження проходження однієї з двох паралельних гіперплощини через точку $\mathbf{x}^{m+1}$ зменшує розмірність задачі на 1 [2]. Тому $d = 2T(m,n-1)$. Підставивши у (13) значення d, отримаємо (12).

Лема 5. *При $n > 1$ виконується нерівність*

$$T(m,n) \le 2^n(m-n+2)\,C_{m-1}^n + 3\sum_{j=0}^{n-1} 2^j C_{m-1}^j. \tag{14}$$

Доведення. Для того, щоб поширити (12) на випадок $n = 1$ покладемо

$$T(m,0) = m+2,\ (m > 0). \tag{15}$$

Використовуючи метод математичної індукції неважко отримати наступне "багатокрокове" узагальнення (12)

$$T(m,n) \leq \sum_{j=0}^{k} 2^j C_k^j T(m-k, n-j), \quad 1 \leq k < n.$$

При $n \leq k < m$ попередня формула ускладнюється і набуває наступного вигляду

$$T(m,n) \leq \sum_{j=0}^{n-1} 2^j C_{m-1}^j T(m-k, n-j) + 2^n \sum_{j=0}^{k-n} C_{m-j-2}^{n-1} T(m-k+j, 0).$$

Підставимо у попередню нерівність $k = m-1$. Тоді з урахуванням (15) і того, що $T(1,n) = 3$, отримаємо

$$T(m,n) \leq 3 \sum_{j=0}^{n-1} 2^j C_{m-1}^j + 2^n \sum_{j=1}^{m-n} C_{m-j-1}^{n-1} T(j, 0). \tag{16}$$

Оцінимо другу суму у попередній нерівності. Для цього використаємо відому комбінаторну рівність

$$C_n^n + C_{n+1}^n + \ldots + C_{n+p-1}^n = C_{n+p}^{n+1}.$$

Тоді

$$\sum_{j=1}^{m-n} C_{m-j-1}^{n-1} T(j,0) \leq T(m-n,0) \sum_{j=1}^{m-n} C_{m-j-1}^{n-1} = (m-n+2)\, C_{m-1}^{n+1}.$$

З останньої нерівності і з (16) отримуємо (14), що й доводить лему.

Використаємо (14) для уточнення верхньої оцінки для $D_2(m,n)$. Оскільки двом різним досяжним 3-розбиттям (A_1, A_2, A_3) і (A_3, A_2, A_1) $(A_1 \cup A_3 \neq \varnothing)$ породженим вектором $\mathbf{w}$ і порогами $t_1,\ t_2$ відповідає єдине д-розбиття $(A_1 \cup A_3, A_2)$ на ДНЕ із структурою $(\mathbf{w}, t_1, t_2)$, то $T(m,n) \geq 2D_2(m,n) - 1$ (неважко показати, що при $m > 1,\ n \geq 1$ має місце сильніша нерівність $T(m,n) > 2D(m,n)$). Тоді з урахуванням (14) при $m > n \geq 1$ отримаємо:

$$D_2(m,n) < 2^{n-1}(m-n+2)\, C_{m-1}^n + \frac{3}{2} \sum_{j=0}^{n-1} 2^j C_{m-1}^j. \tag{17}$$

Якщо порівняти оцінки (17) і (18), то видно, що у коефіцієнт при m^{n+1} у (18) більш, ніж у 8 рази менший, ніж відповідний коефіцієнт у (17) (коефіцієнти перед іншими степенями m у (17) також менші). Тому оцінка (17) є більш точною.

Покажемо, що верхня оцінка для $D_2(m,n)$ може бути використана для доведення наступного факту

Твердження 2. *При $n > 10$ для довільної множини $A \subset \mathbb{R}^n$, такої, що* $\operatorname{Card} A \geq 5n$ *знайдеться розбиття множини A, яке не є д-розбиттям.*

Доведення. Використаємо оцінку (19) та формулу Стірлінга. Отримаємо

$$D_2(5n,n) < 3\cdot\frac{(10n)^{n+1}}{(n+1)!} = \frac{30\cdot 10^n n^{n+1}}{(n+1)\,!} = \frac{30\cdot 10^n n^{n+1} e^{n+1}}{\sqrt{2\pi(n+1)}\,(n+1)^{\ n+1} e^{\theta(n+1)}} < 6\cdot(10e)^n,$$

де $|\theta(n)| < (12n)^{-1}$. Якщо

$$n > \frac{\ln 6}{4\ln 2 - \ln 5e} \approx 10,9822,$$

то $D(5n,n) < 2^{5n}$. Ми отримали, що кількість д-розбиттів множини A менша за кількість усіх її підмножин. Звідси випливає існування розбиттів множини A, які не є д-розбиттями.

Зауваження. За допомогою аналогічних міркувань можна показати існування непорогових 3-розбиттів при $m \geq 2,5n$.

Встановимо тепер обмеження на потужність множини A, яке б забезпечувало д-розділимість довільної дихотомії множини A, тобто виконання умови $D_2(m,n) = 2^m$.

Твердження 3. *Для всіх $n \in \mathbb{N}$ можна вказати таку множину $A_n \subset \mathbb{R}^n$, що* $\operatorname{Card} A_n = 2n$ *і* $D_2(A_n) = 2^{2n}$.

Доведення. Використаємо індукції по розмірності векторного простору, причому будемо доводити, усі розбиття множин A_n, $n \in \mathbb{N}$ можна здійснити на ДНЕ, пороги яких задовольняють умову

$$t_1 = -1, \quad t_2 = 2. \tag{18}$$

Справедливість твердження при $n = 1$ для довільної множини, яка містить дві точки дійсної прямої легко отримати безпосередньою перевіркою (усі розбиття двохелементної множини є лінійно розділимими). Крім того легко переконатися, що усі д-розбиття при $n = 1$ можна отримати за допомогою ДНЕ з порогами, які задовольняють (18). Для визначеності покладемо $A_1 = \{-1, 1\}$. Нехай для всіх $r < n$ твердження вже доведене і доведемо його при . Задамо множину $A_n \subset \mathbb{R}^n$ наступним чином

$$A^n = \left\{(x_1,\ldots,x_{n-1},0) \mid (x_1,\ldots,x_{n-1}) \in A^{n-1}\right\} \cup \{(0,\ldots,0,-1),(0,\ldots,0,1)\}.$$

Покажемо, що усі розбиття множини A_n є д-розбиттями. Нехай (A_n^+, A_n^-) — довільне розбиття множини A_n і нехай $\left(A_{n-1}^+, A_{n-1}^-\right)$ відповідне розбиття вершин множини A_{n-1} $\left(A_{n-1}^+ = \{(x_1,\ldots,x_{n-1}) | \,(x_1,\ldots,x_{n-1},0) \in A_n^+\},\, A_{n-1}^- = A_{n-1} \backslash A_{n-1}^+\right)$. За припущенням індукції його можна здійснити за допомогою ДНЕ із структурою $((w_1,\ldots,w_{n-1}),-1,2)$. Можливими є 4 випадки:

1) $\{(0,\ldots,0,-1),\ (0,\ldots,0,1)\} \subset A_n^+$. У цьому випадку покладемо $w_n = 3$;

2) $(0,\ldots,0,-1) \in A_n^+$, $(0,\ldots,0,1) \in A_n^-$. У цьому випадку покладемо $w_n = 1,5$;

3) $(0,\ldots,0,-1) \in A_n^-$, $(0,\ldots,0,1) \in A_n^+$. У цьому випадку покладемо $w_n = -1,5$;

4) $\{(0,\ldots,0,-1)\,,\ (0,\ldots,0,1)\} \subset A_n^-$. У цьому випадку покладемо $w_n = 0$.

Легко переконатися, що у кожного з чотирьох випадків ДНЕ із структурою $((w_1,\ldots,w_{n-1},w_n)\,,-1,2)$ здійснює д-розбиття множини A_n, причому пороги t_1, t_2 задовольняють (18). Твердження доведене.

Зауваження. Цікаво співставити твердження 3 та відому теорему Радона [14], згідно якої для довільної множини $A \subset \mathbb{R}^n$, для якої $\operatorname{Card} A > n+1$, знайдеться така дихотомія (A^+, A^-), яка не є лінійно розділимою.

Зауваження. На основі тверджень 2 і 3 можна зробити висновок, що для класу $\mathcal{F}$ ДНЕ з дійсними вагами $2n \le \operatorname{VCDim} \mathcal{F} \le 5n$. Використовуючи методи, наведені у [2] при дослідженні нейромереж прямого поширення, побудованих із поліноміальних нейронних елементів, та той факт, що д-розбиття можна реалізувати за допомогою двошарової нейромережі із звичайних порогових елементів, можна показати, що для класу $\mathcal{FN}$ нейромереж прямого поширення, побудованих із ДНЕ $\operatorname{VCDim} \mathcal{FN} = O(m \log m)$, де m — кількість вільних параметрів мережі (кількість вагових коефіцієнтів та порогів усіх нейронів мережі).

1. *Haykin S.* Neural Networks. A Comprehensive Foundation. — N.Y.: Prentice Hall, Inc. 1999. – 804 c.
2. *M. H. Hassoun* Fundamentals of Artificial Neural Networks. — MIT Press, 1995. – 410 c.
3. *R. Takiyama.* Multiple threshold perceptron, Pattern recognition., vol. 10, pp. 27-30, 1978.
4. *R. Takiyama.* The separating capacity of multi-threshold threshold element, IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-7, pp. 112-116, Jan. 1985.
5. *S. Olafsson and Y. S. Abu-Mostafa.* The capacity of multilevel threshold function, IEEE Trans. Pattern Anal. Machine Intell., vol. 10, NO. 2, pp. 277-281, March 1988.
6. *Шоломов Л. А.* О реализаций булевой функции на многопороговом елементе. — Автоматика и телемеханика, 1966, №3, С. 97–104.
7. *Бахарев А. Т.* Оптимизация многопороговых моделей. — Пробл. случайного поиска (Рига), 1975, вып. 4, С. 209–214.
8. *D. Haring.* Multi-threshold threshold elements, IEEE Trans. Electron. Comput., vol. EC-15, pp. 45-65, June 1965.
9. *V. Deolalikar.* A Two-Layer Paradigme Capable of Forming Arbitrary Decision Regions in Input Space, IEEE Trans. on Neural Networks, vol. TNN-4, No. 2 pp. 343-347, March 2001.
10. *Гече Ф. Е.* Реалізація бульових функцій на двопорогових нейронних елементах // Наук. вісник Ужгород. ун-ту. Сер. матем. – Ужгород, 1999. — Вип. 4 С. – 17-24.
11. *Ф. Гече, А. Батюк, В. Коцовський.* Властивості бульових функцій реалізовних на двопорогових елементах // Вісник Національного університету "Львівська політехніка". Комп'ютерна інженерія та інформаційні технології. – Львів, 2001. – №438. – С. 22-25.
12. *Руденко О. Г., Бодянський Є. В.* Штучні нейронні мережі. — Харків: ТОВ "Компанія СМІТ", 2006. — 404 с.
13. *В. Н. Вапник, А. Я. Червоненкис.* Теория распознавания образов. — М.: Наука – 1974. — 416 с.
14. *К. Лейхтвейс.* Выпуклые множества. — М.: Наука – 1985. — 335 с.