Угол между плоскостями линейных тестовых алгоритмов

М. В. Носов[1]

В работе исследуется вопрос об угле между нормальными векторами разделяющих гиперплоскостей, задаваемых линейными тестовыми алгоритмами. Вопрос рассматривается для различных опорных множеств тестов.

Ключевые слова: линейные тестовые алгоритмы, разделяющие гиперплоскости, тупиковые тесты.

В теории тестовых алгоритмов рапознавания выделяются три линейных алгоритма, исследуем вопрос о величине угла между соответствующими гиперплоскостями. Пусть E^n - n-мерный единичный куб, (T_0, T_1) - материал обучения, $T_0, T_1 \subset E^n, T_0 \bigcap T_1 = \emptyset, T_0 \neq \emptyset, T_1 \neq \emptyset$, пусть T - опорное множество тестов.

Первый алгоритм - линейный тестовый алгоритм[1][3], задаваемый решающим правилом $R_1(t), t = (t_1, \ldots, t_n)$

$$R_1(t) = \sum_{i=1}^{n} p_i t_i + p_0,$$

$$p_i = \tfrac{1}{|T|} \sum_{\tau \in T} \tau_i, i = 1, \ldots, n,$$

$p = (p_1, \ldots, p_n)$ - вектор информационных весов,

p_0 - порог - действительное число.

Второй алгоритм - линейный тестовый алгоритм Журавлёва Ю.И.[2], решающее правило имеет вид

$$R_2(t) = 2\sum_{i=1}^{n} p_i \Delta s_i t_i - \sum_{i=1}^{n} p_i \Delta s_i,$$

$$\Delta s_i = s_{1i} - s_{0i},$$

$$s_{1i} = \tfrac{1}{|T_1|} \sum_{\theta \in T_1} \theta_i \text{ - центр тяжести } T_1,$$

$$s_{0i} = \tfrac{1}{|T_0|} \sum_{\theta \in T_0} \theta_i \text{ - центр тяжести } T_0,$$

[1] *Носов Михаил Васильевич* — с.н.с. каф. математической теории интеллектуальных систем мех.-мат. ф-та МГУ, e-mail: mvnosov@rambler.ru.

Nosov Michail Vasilevich-senior researcher, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Mathematical Theory of Intellectual Systems.

$$\theta = (\theta_1, \ldots, \theta_n),$$
$$\Delta s_i = s_{1i} - s_{0i},$$
$$i = 1, \ldots, n,$$
$$\Delta s = (\Delta s_1, \ldots, \Delta s_n).$$

Третий тестовый алгоритм - линейный тестовый алгоритм - наилучшее линейное приближение алгоритма голосования Кудрявцева В.Б.[1][3], его решающее правило имеет вид

$$R_3(t) = 4\sum_{i=1}^{n} q_i \Delta s_i t_i - 2\sum_{i=1}^{n} q_i \Delta s_i,$$

$$q_i = \sum_{\tau \in T} \frac{\tau_i}{2^{|\tau|}}, i = 1, \ldots, n.$$

Таким образом, имеется три нормальных вектора

$$\omega_1 = (p_1, \ldots, p_n),$$
$$\omega_2 = (p_1 \Delta s_1, \ldots, p_n \Delta s_n),$$
$$\omega_3 = (q_1 \Delta s_1, \ldots, q_n \Delta s_n).$$

Оценим угол между этими векторами для некоторых случаев материала обучения (T_0, T_1) и опорного множества тестов T.

1.Пусть $T_0 = \{(0, \ldots, 0)\}, T_1 = E^{n,k}$ - k-ый слой куба, $1 \leq k \leq n$, в качестве опорного множества возьмём все тупиковые тесты, очевидно, $T = E^{n,n-k+1}$. Ввиду "равноправия"всех координат все три вектора сонаправлены.

2.Пусть n - чётное, $n = 2l$. Пусть T_0 состоит из одного вектора, у которого первые l координат 1, остальные 0:

$$T_0 = \{(1, \ldots, 1, 0, \ldots, 0)\},$$

условно можно написать

$$T_0 = \{(0, \ldots, 0)\} \oplus (1, \ldots, 1, 0, \ldots, 0).$$

T_1 представить в виде

$$T_1 = E^{n,k} \oplus (1, \ldots, 1, 0, \ldots, 0).$$

Очевидно, что $T = E^{n,n-k+1}$ и вектор p сонаправлен единичному. При $1 \leq i \leq l, s_{0i} = 1$,

$$s_{1i} = \frac{1}{\binom{n}{k}} \sum_{\theta \in T_1} \theta_i = \frac{1}{\binom{n}{k}} \sum_{\theta \in E^{n,k}} (\theta_i \oplus 1) = \frac{1}{\binom{n}{k}} \sum_{\theta \in E^{n,k}} (1 - \theta_i) =$$

$$= 1 - \frac{1}{\binom{n}{k}} \binom{n-1}{k-1} = 1 - \frac{k}{n(n-k+1)},$$

$$\Delta s_i = -\frac{k}{n(n-k+1)}.$$

При $l \leq i \leq 2l, s_{0i} = 0$,

$$s_{1i} = \frac{k}{n(n-k+1)},$$

$$\Delta s_i = \frac{k}{n(n-k+1)}.$$

Следовательно, у вектора Δs все координаты одинаковы по абсолютной величине, но в первой половине отрицательны, а во второй положительны. Значит вектор w_1 перпендикулярен вектору w_2 и вектору w_3.

3.Перпендикулярность векторов в предыдущем случае основывалась на положительности половины координат и отрицательности другой половины для вектора, соединяющего центры тяжестей T_0 и T_1. Всегда можно поменять в признаке 0 и 1, это не приведёт к изменению опорного множества, но все координаты вектора Δs будут одного знака. Итак, $\Delta s \geq 0$; ясно, что $1/|T| \leq p_i \leq 1$ для ненулевых p_i и пусть хотя бы в одной координате вектора p и Δs положительны.Тогда имеем

$$\cos(\widehat{w_1, w_2}) = \frac{\sum\limits_{i=1}^{n} p_i^2 \Delta s_i}{\sqrt{\sum\limits_{i=1}^{n} p_i^2 \Delta s_i^2} \sqrt{\sum\limits_{i=1}^{n} p_i^2}} \geq \frac{\sum\limits_{i=1}^{n} p_i^2 \Delta s_i}{\sqrt{\sum\limits_{i=1, p_i \neq 0} \Delta s_i^2 \sqrt{n}}} \geq$$

$$\geq \frac{\frac{1}{|T|^2} \sum\limits_{i=1, p_i \neq 0} \Delta s_i}{\sqrt{\sum\limits_{i=1, p_i \neq 0} \Delta s_i^2 \sqrt{n}}} \geq \frac{1}{|T|^2 \sqrt{n}},$$

$$(\widehat{w_1, w_2}) \leq \arccos \frac{1}{|T|^2 \sqrt{n}}. \tag{1}$$

Правую часть неравенства можно увеличить, доведя до соотношения от n

$$(\widehat{w_1, w_2}) \leq \arccos \frac{1}{(\binom{n}{[n/2]})^2 \sqrt{n}}.$$

Далее, при $p_i > 0$

$$\frac{|T|}{2^{\max|\tau|}} p_i \leq \sum_{\tau \in T} \frac{\tau_i}{2^{|\tau|}} \leq \frac{|T|}{2^{\min|\tau|}} p_i,$$

$$\cos(\widehat{w_1, w_3}) = \frac{\sum\limits_{i=1}^{n} p_i q_i \Delta s_i}{\sqrt{\sum\limits_{i=1}^{n} q_i^2 \Delta s_i^2}\sqrt{\sum\limits_{i=1}^{n} p_i^2}} \geq \frac{\frac{|T|}{2^{\max|\tau|}}\sum\limits_{i=1}^{n} p_i^2 \Delta s_i}{\frac{|T|}{2^{\min|\tau|}}\sqrt{\sum\limits_{i=1}^{n} p_i^2 \Delta s_i^2}\sqrt{\sum\limits_{i=1}^{n} p_i^2}} \geq$$

$$\geq \frac{1}{2^{\max|\tau|-\min|\tau|}|T|^2\sqrt{n}},$$

$$(\widehat{w_1, w_3}) \leq \arccos \frac{1}{2^{\max|\tau|-\min|\tau|}|T|^2\sqrt{n}}. \tag{2}$$

Доводим соотношение до n

$$(\widehat{w_1, w_3}) \leq \arccos \frac{1}{2^{n-1}(\binom{n}{[n/2]})^2\sqrt{n}}.$$

Пример. Пусть $T_0 = \{(0, \dots, 0)\}, T_1 = \{(1, 1, 0, 0, \dots, 0), (1, 0, 1, 0, \dots, 0), \dots, (1, 0, 0, 0, \dots, 1)\}$. Очевидно, множество тупиковых тестов $T = \{(1, 0, \dots, 0), (0, 1, \dots, 1)\}$. Тогда

$$w_1 = p = (\frac{1}{2}, \dots, \frac{1}{2}), q = (\frac{1}{2}, \frac{1}{2^{n-1}}, \dots, \frac{1}{2^{n-1}}),$$

$$s_1 = (1, \frac{1}{n-1}, \frac{1}{n-1}, \dots, \frac{1}{n-1}), s_0 = (0, \dots, 0),$$

$$\Delta s = (1, \frac{1}{n-1}, \frac{1}{n-1}, \dots, \frac{1}{n-1}),$$

$$w_2 = p\Delta s = (\frac{1}{2}, \frac{1}{2(n-1)}, \frac{1}{2(n-1)}, \dots, \frac{1}{2(n-1)}),$$

$$w_3 = q\Delta s = (\frac{1}{2}, \frac{1}{2^{n-1}(n-1)}, \frac{1}{2^{n-1}(n-1)}, \dots, \frac{1}{2^{n-1}(n-1)}).$$

Из неравенства (1) следует

$$(\widehat{w_1, w_2}) \leq \arccos \frac{1}{4\sqrt{n}},$$

Точное значение

$$(\widehat{w_1, w_2}) = \arccos \frac{2}{\sqrt{n}\sqrt{1 + \frac{1}{n-1}}}.$$

Из неравенства (2) следует

$$(\widehat{w_1, w_3}) \leq \arccos \frac{1}{2^n\sqrt{n}},$$

Точное значение

$$(\widehat{w_1, w_3}) = \arccos \frac{1 + \frac{1}{2^{n-1}}}{\sqrt{n}\sqrt{1 + \frac{1}{2^{2n-4}(n-1)}}}.$$

4.Исследуем угол между w_2 и w_3 в общем случае

$$\cos(\widehat{w_2, w_3}) = \frac{\sum\limits_{i=1}^{n} p_i q_i \Delta s_i^2}{\sqrt{\sum\limits_{i=1}^{n} q_i^2 \Delta s_i^2}\sqrt{\sum\limits_{i=1}^{n} p_i^2 \Delta s_i^2}} \geq \frac{\frac{|T|}{2^{\max|\tau|}} \sum\limits_{i=1}^{n} p_i^2 \Delta s_i^2}{\frac{|T|}{2^{\min|\tau|}} \sqrt{\sum\limits_{i=1}^{n} p_i^2 \Delta s_i^2}\sqrt{\sum\limits_{i=1}^{n} p_i^2 \Delta s_i^2}} =$$

$$= \frac{1}{2^{\max|\tau| - \min|\tau|}},$$

$$(\widehat{w_2, w_3}) \leq \arccos \frac{1}{2^{\max|\tau| - \min|\tau|}}.$$

Доводим соотношение до n

$$(\widehat{w_2, w_3}) \leq \arccos \frac{1}{2^{n-1}}.$$

5. Используем результат Коршунова А.Д. для класса монотонных функций $M(n)$.

Лемма [4]. У почти всех функций из $M(n)$ нижние единицы расположены в $E^{n,n/2-1}, E^{n,n/2}, E^{n,n/2+1}$ при чётном n и в $E^{n,(n-3)/2}, E^{n,(n-1)/2}, E^{n,(n+1)/2}, E^{n,(n+3)/2}$ при нечётном n.

Из выражения $\cos(\widehat{w_2, w_3})$ видно, что можно считать все Δs_i неотрицательными.

Рассмотрим чётный случай, $n = 2l$. В качестве T берём тупиковые тесты, лежащие в $E^{n,n/2-1}, E^{n,n/2}, E^{n,n/2+1}$. Введём обозначения

$$a_i = |\{\tau \in T | \tau_i = 1, |\tau| = l - 1\}|,$$
$$b_i = |\{\tau \in T | \tau_i = 1, |\tau| = l\}|,$$
$$c_i = |\{\tau \in T | \tau_i = 1, |\tau| = l + 1\}|.$$

Тогда

$$|T|p = (\ldots, a_i + b_i + c_i, \ldots), 2^l q = (\ldots, 2a_i + b_i + \frac{c_i}{2}, \ldots),$$

$$2^l q = |T|p + z, z = (\ldots, a_i - \frac{c_i}{2}, \ldots).$$

Имеет место следующее неравенство

$$-\frac{1}{2}|T|p_i \le z_i \le |T|p_i, i = 1, \dots, n.$$

Угол между векторами w_2 и w_3 равен углу между векторами u_2 и u_3

$$u_2 = (p_1|T|\Delta s_1, \dots, p_n|T|\Delta s_n),$$

$$u_3 = (2^l q_1 \Delta s_1, \dots, 2^l q_n \Delta s_n).$$

Далее

$$u_3 = u_2 + (z_1|\Delta s_1|, \dots, z_n|\Delta s_n|),$$

как было показано

$$-\frac{1}{2}|T|p_i|\Delta s_i| \le z_i|\Delta s_i| \le |T|p_i|\Delta s_i|,$$

таким образом, угол между векторами u_2 u_3 не превышает угла между вектором u_2 и вектором с концом в вершине параллелепипеда

$$\prod_{i=1}^{n}[-\frac{1}{2}|T|p_i|\Delta s_i|, |T|p_i|\Delta s_i|] + u_2 =$$

$$= \prod_{i=1}^{n}[\frac{1}{2}|T|p_i|\Delta s_i|, 2|T|p_i|\Delta s_i|].$$

Достаточно очевидно, что максимум величины угла достигается на одной из вершин $(2^{g_1}|T|_1\Delta s_1, \dots, 2^{g_n}|T|_1\Delta s_n)$, пусть $g_i = -1, i \in J_1, g_i = 1, i \in J_2, J_1, J_2 \subset \{1, \dots, n\}$. Проведём через эту вершину гиперплоскость с направляющим вектором u_2, её уравнение имеет вид

$$\sum_{i=1}^{n} p_i|T||\Delta s_i|t_i - \frac{1}{2}\sum_{i\in J_1} p_i^2|T|^2|\Delta s_i|^2 - 2\sum_{i\in J_2} p_i^2|T|^2|\Delta s_i|^2 = 0,$$

эта гиперплоскость пересекает прямую, проходящую через начало координат, с направляющим вектором u_2 в точке λu_2, где λ удовлетворяет уравнению

$$\sum_{i=1}^{n} p_i^2|T|^2|\Delta s_i|^2\lambda - \frac{1}{2}\sum_{i\in J_1} p_i^2|T|^2|\Delta s_i|^2 - 2\sum_{i\in J_2} p_i^2|T|^2|\Delta s_i|^2 = 0,$$

Пусть α - угол между вектором u_2 и вектором с началом в начале координат и концом в точке $(2^{g_1}|T|_1\Delta s_1, \dots, 2^{g_n}|T|_1\Delta s_n)$, тогда

$$\tan^2\alpha = \frac{\sum\limits_{i\in J_1}(\frac{1}{2}-\lambda)^2 p_i^2|T|^2|\Delta s_i|^2 + \sum\limits_{i\in J_2}(2-\lambda)^2 p_i^2|T|^2|\Delta s_i|^2}{\lambda^2\sum\limits_{i=1}^{n} p_i^2|T|^2|\Delta s_i|^2}.$$

Пусть

$$P_1 = \sum_{i \in J_1} p_i^2 |T|^2 |\Delta s_i|^2, P_2 = \sum_{i \in J_2} p_i^2 |T|^2 |\Delta s_i|^2,$$

Тогда

$$\tan^2 \alpha = \frac{(\frac{1}{2} - \lambda)^2 P_1 + (2 - \lambda)^2 P_2}{\lambda^2 (P_1 + P_2)},$$

$$\lambda = \frac{\frac{1}{2} P_1 + 2 P_2}{P_1 + P_2}.$$

Несложно найти максимальное значение $\tan \alpha$ равное $\frac{3}{4}$. Таким образом, при чётном n и векторах опорного множества, лежащих во множествах $E^{n,n/2-1}, E^{n,n/2}, E^{n,n/2+1}$, угол между векторами w_2 и w_3 не превышает $\arctan \frac{3}{4}$, $(\arctan \frac{3}{4} = 36,87^\circ)$.

Рассмотрим нечётный случай, $n = 2l+1$ аналогично чётному. В качестве T берём тупиковые тесты, лежащие в $E^{n,(n-3)/2}, E^{n,(n-1)/2}, E^{n,(n+1)/2}, E^{n,(n+3)/2}$. Введём обозначения

$$a_i = |\{\tau \in T | \tau_i = 1, |\tau| = l - 1\}|,$$
$$b_i = |\{\tau \in T | \tau_i = 1, |\tau| = l\}|,$$
$$c_i = |\{\tau \in T | \tau_i = 1, |\tau| = l + 1\}|,$$
$$d_i = |\{\tau \in T | \tau_i = 1, |\tau| = l + 2\}|.$$

Тогда

$$|T| p = (\ldots, a_i + b_i + c_i + d_i, \ldots),$$
$$2^{l+\frac{1}{2}} q = (\ldots, 2^{\frac{3}{2}} a_i + 2^{\frac{1}{2}} b_i + 2^{-\frac{1}{2}} c_i + 2^{-\frac{3}{2}} d_i, \ldots),$$
$$2^{l+\frac{1}{2}} q = |T| p + z,$$
$$z = (\ldots, (2\sqrt{2} - 1) a_i + (\sqrt{2} - 1) b_i - (1 - \frac{1}{\sqrt{2}}) c_i - (1 - \frac{1}{2\sqrt{2}}) d_i, \ldots).$$

Имеет место следующее неравенство

$$-\frac{2\sqrt{2} - 1}{2\sqrt{2}} |T| p_i \leq z_i \leq (2\sqrt{2} - 1) |T| p_i, i = 1, \ldots, n.$$

Угол между векторами w_2 и w_3 равен углу между векторами u_2 и u_3

$$u_2 = (p_1 |T| \Delta s_1, \ldots, p_n |T| \Delta s_n),$$
$$u_3 = (2^{l+\frac{1}{2}} q_1 \Delta s_1, \ldots, 2^{l+\frac{1}{2}} q_n \Delta s_n).$$

Далее производя вычисления, как в чётном случае приходим к результату.При нечётном n и векторах опорного множества, лежащих во множествах $E^{n,(n-3)/2}, E^{n,(n-1)/2}, E^{n,(n+1)/2}, E^{n,(n+3)/2}$, угол между векторами w_2 и w_3 не превышает $\arctan \frac{\sqrt{7\sqrt{2}}}{4}$, $(\arctan \frac{\sqrt{7\sqrt{2}}}{4} = 38,19^\circ)$.

Список литературы

[1] Константинов Р.М., Королёва З.Е., Кудрявцев В.Б. Комбинаторно-логический подход к задачам прогноза рудоносности Проблемы кибернетики 1976 31 5–33

[2] Дмитриев А.Н., Журавлёв Ю.И., Кренделёв Ф.П. О математических принципах классификации предметов и явлений Дискретный анализ 1966 7 3–15

[3] Алешин С.В. Распознавание динамических образов Издательство Московского университета Москва 1996 97

[4] Коршунов А.Д. О числе монотонных булевых функций Проблемы кибернетики 1981 38 5–108

The angle between the planes of linear test algorithms

Nosov M.V.

The paper investigates the question of the angle between the normal vectors of separating hyperplanes defined by linear test algorithms. The question is considered for various reference sets of tests.

Keywords: linear test algorithms separating hyperplanes, dead-end tests.

References

[1] Konstantinov R.M., Koroleva Z.E., Kudryavtsev V.B., “Combinatorial-logical approach to the problems of ore-bearing prediction”, *Problems of Cybernetics*, **31** (1976), 5–33 (In Russian).

[2] Dmitriev A.N., Zhuravlev Yu.I., Krendelev F.P., “On mathematical principles of classification of objects and phenomena”, *Discrete analysis*, **7** (1966), 3–15 (In Russian).

[3] Aleshin S.V., *Dynamic image recognition*, Moscow University Press, Moscow, 1996 (In Russian), 97 c.

[4] Korshunov A.D., “On the number of monotone Boolean functions”, *Problems of Cybernetics*, **38** (1981), 5–108 (In Russian)