Приближенные процедуры голосования*

М.В. Носов

В работе формулируются две процедуры распознавания, основой которых явился алгоритм голосования по опорному множеству тестов. Исследуются их решающие правила.

Пусть E_2^n – n-мерный единичный куб, (T_1, T_0) – материал обучения, $\mathcal{M}$ – опорное множество тестов. Разрешим алгоритму голосования по опорному множеству тестов «ошибаться» в некоторых местах [1], [2], то есть, формально, пусть m – целое число, $0 \leqslant m \leqslant n$, вектор $\theta, \theta \in T_j, j = 0, 1$ голосует по тесту τ за принадлежность распознаваемого вектора t к j-му множеству, если t отлично от θ не более чем m координатах, в которых у вектора τ стоят 1. Решающее правило можно записать в следующем виде:

$$G(m,t) = G_1(m,t) - G_0(m,t), \text{ где}$$

$$G_j(m,t) = \frac{1}{m_j} \sum_{\tau \in \mathcal{M}} \sum_{\theta \in T_j} g_m(t,\tau,\theta), m_j = |T_j|, j = 0, 1,$$

$$g_m(t,\tau,\theta) = \begin{cases} 1, & \text{если } |\{i | t_i \neq \theta_i, \text{ при } \tau_i = 1\}| \leqslant m, \\ 0, & \text{иначе.} \end{cases}$$

---

*Работа выполнена при поддержке гранта РФФИ №99-01-00317

Лемма 1. *Имеет место следующее равенство:*

$$g_m(t,\tau,\theta) = 1 - (-1)^{m+1}\prod_{k=0}^{m}\Big(\sum_{\substack{\beta \\ \beta\subset\{1,\dots,n\} \\ |\beta|=n-k}}\Big(\prod_{i\in\beta}(1-\tau_i|t_i-\theta_i|)\Big)-1\Big).$$

Доказательство. Действительно, если вектора t и θ отличаются больше чем в m координатах, в которых у вектора τ стоят единицы, тогда обязательно для любого подмножества β выполняется

$$\sum_{\substack{\beta \\ |\beta|=n-k}}\Big(\prod_{i\in\beta}(1-\tau_i|t_i-\theta_i|)\Big)-1=-1, 0\leqslant k\leqslant m.$$

Если вектора t и θ отличаются в r координатах и $r\leqslant m$, то

$$\sum_{\substack{\beta \\ |\beta|=n-r}}\Big(\prod_{i\in\beta}(1-\tau_i|t_i-\theta_i|)\Big)=1,$$

так как только одно слагаемое в левой части равно 1 (при β равном теоретико-множественному дополнению к этим r координатам), а все остальные слагаемые равны 0. Лемма доказана.

Исследуем выражение

$$h_m(x) = 1-(-1)^{m+1}\prod_{k=0}^{m}\Big(\sum_{\substack{\beta \\ \beta\subset\{1,\dots,n\} \\ |\beta|=n-k}}\prod_{i\in\beta}(1-x_i)-1\Big),$$

$$x=(x_1,x_2,\dots,x_n).$$

Представим его в виде

$$h_m(x) = \sum_{\substack{\beta \\ \beta\subset\{1,\dots,n\}}} a_\beta\prod_{i\in\beta}(x_i).$$

Очевидно, что

$$h_m(x_0)=\begin{cases}1, & \text{если } |x_0|\leqslant m,\\ 0, & \text{если } |x_0|\geqslant m+1.\end{cases}$$

Значит, $a_\emptyset=1$ и $a_\beta=0$, если $1\leqslant|\beta|\leqslant m$. Из симметричности $h_m(x)$ по индексам следует, что $a_\beta=a_\gamma$ при $|\beta|=|\gamma|$, положим $a_\beta=b_l$ при $|\beta|=l$, тогда можно записать

$$b_{m+i}+\mathrm{C}^1_{m+i}b_{m+i-1}+\mathrm{C}^2_{m+i}b_{m+i-2}+\ldots+\mathrm{C}^{i-1}_{m+i}b_{m+1}+1=0.$$

Решением является $b_{m+i}=(-1)^i\mathrm{C}^{i-1}_{m+i-1}=(-1)^i\mathrm{C}^m_{m+i-1}$.

Предложение 1. *Решающим правилом алгоритма голосования при наличии не более m различий в распознаваемом и голосующем векторе является многочлен*

$$G(m,t)=G_1(m,t)-G_0(m,t), \text{ где}$$

$$G_j(m,t)=(-1)^m\sum_{\substack{\beta \\ \beta\subset\{1,\dots,n\} \\ |\beta|\geqslant m+1}}(-1)^{|\beta|}\mathrm{C}^m_{|\beta|-1}\Big(\sum_{\tau\in\mathcal{M}}\prod_{i\in\beta}\tau_i\Big)\Big(\frac{1}{m_j}\sum_{\theta\in T_j}\prod_{i\in\beta}|t_i-\theta_i|\Big),$$

$$j=0,1.$$

Многочлены $G_j(m,t)$ отличаются от соответствующих многочленов голосования наличием весовых коэффициентов при слагаемых по множествам β одинаковой мощности.

В качестве еще одной модификации алгоритма голосования рассмотрим следующий алгоритм. Пара векторов $(u,v), u\in T_1, v\in T_0$ голосует за принадлежность вектора t таблице $T_1(T_0)$ по тесту τ, если во всех координатах, в которых у вектора τ стоят единицы и вектора u и v отличаются в этих координатах, вектор t совпадает с вектором $u(v)$. Решающее правило имеет вид

$$\Lambda(t)=\Lambda_1(t)-\Lambda_0(t),$$

$$\Lambda_1(t) = \frac{1}{m_0 m_1} \sum_{\tau\in\mathcal{M}} \sum_{u\in T_1} \sum_{v\in T_0} \prod_{i=1}^{n} (1-\tau_i|t_i-u_i||u_i-v_i|),$$

аналогично $\Lambda_0(t)$. Преобразуем $\Lambda_1(t)$:

$$\Lambda_1(t) = \sum_{\substack{\beta \\ \beta\subset\{1,\ldots,n\}}} (-1)^{|\beta|} \Big(\sum_{\tau\in\mathcal{M}} \prod_{i\in\beta} \tau_i\Big)\Big(\frac{1}{m_1} \sum_{u\in T_1} \prod_{i\in\beta} |t_i-u_i| \Big(\frac{1}{m_0} \sum_{v\in T_0} |u_i-v_i|\Big)\Big).$$

Возьмем линейное приближение $\Lambda(t)$ относительно величин $|t_i-u_i|$ и $|t_i-v_i|$, $i=1,\ldots,n$

$$\Lambda^{(1)}(t) = \sum_{i=1}^{n} p_i(S_{0i}-S_{1i}) + 2\sum_{i=1}^{n} p_i(S_{1i}-S_{0i})t_i,$$

где $p=(p_1,\ldots,p_n)$ – вектор информационных весов $\mathcal{M}$, $S_j = (S_{j1},\ldots,S_{jn})$ – спектр T_j, $j=0,1$.

Предложение 2. *Линейное приближение алгоритма голосования с решающим правилом $\Lambda(t)$ совпадает с аналогичным линейным приближением алгоритма голосования.*

Список литературы

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

[2] Алешин С.В. Распознавание динамических образов. М.: Изд-во МГУ, 1996.

Дифференциальная компьютерная алгебра*

Е.В. Панкратьев

Компьютерная алгебра имеет дело с конструктивными методами в алгебраических структурах и их реализацией в виде программ для ЭВМ. Своим возникновением и развитием компьютерная алгебра обязана бурному развитию вычислительной техники и применению компьютеров в алгебраических исследованиях. Как отдельное направление компьютерная алгебра сформировалась во 2-й половине XX века. В течение достаточно долгого времени для алгоритмов компьютерной алгебры использовались такие термины, как «аналитические вычисления», «символьные вычисления», «аналитические преобразования» и т.д. В настоящее время имеется достаточно обширная литература по компьютерной алгебре, из которой выделим книги [9], [6], [3], [25].

Дифференциальная алгебра как область математики сформировалась в работах Дж. Ритта в первой половине XX столетия. Громадную роль в ее развитии сыграли работы Е. Колчина. И. Капланский в авторском предисловии к книге [8] пишет «Дифференциальную алгебру легко описать: она состоит (на 99% или более) из работ Ритта и Колчина». За полвека, прошедшие с момента публикации книги Капланского, ситуация существенно изменилась, хотя монографии [47] и [34] остаются самыми фундаментальными трудами по дифференциальной алгебре. Важным событием в дифференциальной алгебре стал доклад Колчина [33] на Московском Математическом Конгрессе в 1966 г. В нем Колчин фактически сформулировал программу

*Работа выполнена при финансовой поддержке INTAS, грант 99-1222.