ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 10 за 2017 г.

ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

Яшунский А.Д.

Конечные системы операций для аппроксимации дискретных вероятностных распределений

Рекомендуемая форма библиографической ссылки: Яшунский А.Д. Конечные системы операций для аппроксимации дискретных вероятностных распределений // Препринты ИПМ им. М.В.Келдыша. 2017. № 10. 7 с. doi:10.20948/prepr-2017-10
URL: http://library.keldysh.ru/preprint.asp?id=2017-10

О р д е н а   Л е н и н а
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В. Келдыша
Р о с с и й с к о й   а к а д е м и и   н а у к

А. Д. Яшунский

Конечные системы операций для аппроксимации дискретных вероятностных распределений

Москва — 2017

Яшунский А. Д.

Конечные системы операций для аппроксимации дискретных вероятностных распределений

Рассматриваются распределения случайных величин над конечным множеством, получаемых с помощью конечного набора операций из независимых случайных величин, имеющих заданные распределения. Показано, что для любого конечного множества существует конечный набор операций, позволяющих аппроксимировать любое наперед заданное распределение.

Ключевые слова: случайная величина, конечное множество, аппроксимация, конечная система операций

Alexey Dmitrievich Yashunsky

Finite systems of operations for discrete probability distributions approximation

We consider distributions of random variables over a finite set, obtained as results of operations from a finite set on independent random variables with given distributions. We show that for any finite set there exists a finite set of operations allowing approximation of an arbitrary given distribution.

Key words: random variable, finite set, approximation, finite set of operations

Работа выполнена при поддержке программы фундаментальных исследований ОМН РАН «Алгебраические и комбинаторные методы математической кибернетики и информационные системы нового поколения» (проект «Задачи оптимального синтеза управляющих систем»).

Оглавление

Алгебры вероятностных распределений

Напомним, что пара $\mathfrak{A} = \langle A, \Omega \rangle$ называется *алгеброй*, если Ω — некоторое множество операций, определенных на множестве A, см. [1]. Если $\Omega = \{F_0, F_1, \ldots\}$, то также будем писать $\mathfrak{A} = \langle A; F_0, F_1, \ldots \rangle$. Множество A называется *основным множеством* алгебры. Далее рассматриваем алгебры, у которых основное множество конечно; их называют конечными алгебрами. Без ограничения общности можно считать, что основное множество конечной алгебры есть $E_k = \{0, 1, \ldots, k-1\}$ для некоторого значения k.

Рассмотрим случайные величины со значениями в множестве E_k. Распределение $\mathbf{p}$ такой случайной величины — вектор с k компонентами $\mathbf{p} = (p_0, \ldots, p_{k-1})$, лежащий в $(k-1)$-мерном симплексе, заданном соотношениями $\sum p_i = 1$, $p_i \geqslant 0$, $i = 0, \ldots, k-1$. Соответствующий симплекс будем далее обозначать $T^{(k)}$. *Носителем* распределения $\mathbf{p}$ будем называть множество $N(\mathbf{p}) = \{i \in E_k \mid p_i > 0\}$.

Пусть задана некоторая конечная алгебра $\langle E_k, B \rangle$, где $B \subseteq P_k$. Если $f(x_1, \ldots, x_n) \in B$ и $X_1, \ldots, X_n$ — независимые в совокупности случайные величины, имеющие распределения $\mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(n)}$ соответственно, то $f(X_1, \ldots, X_n)$ есть случайная величина X с распределением $\mathbf{p}$, для компонент которого выполнены равенства:

$$p_i = \sum_{\substack{(\sigma_1, \ldots, \sigma_n) : \\ f(\sigma_1, \ldots, \sigma_n) = i}} p^{(1)}_{\sigma_1} \cdots p^{(n)}_{\sigma_n}.$$

Таким образом, каждая n-арная функция $f \in B$ индуцирует полилинейное отображение $\hat{f} : (T^{(k)})^n \to T^{(k)}$. Обозначим $\hat{B} = \{\hat{f} \mid f \in B\}$. Тогда $\langle T^{(k)}, \hat{B} \rangle$ есть алгебра вероятностных распределений, индуцированная алгеброй $\langle E_k, B \rangle$.

Пусть $G \subset T^{(k)}$ и пусть $W_B(G)$ — наименьшее по включению топологически замкнутое (относительно топологии, заданной стандартной евклидовой метрикой) множество, содержащее G и замкнутое относительно всех операций $\hat{f}, f \in B$. Тогда $\langle W_B(G), \hat{B} \rangle$ — подалгебра алгебры $\langle T^{(k)}, \hat{B} \rangle$, которую будем называть *алгеброй аппроксимируемых распределений, порожденной начальным множеством G и операциями B*. В случае, когда множество G состоит из единственного распределения $\mathbf{p}$, вместо $W_B(\{\mathbf{p}\})$ будем писать $W_B(\mathbf{p})$.

Имеют место следующие почти очевидные свойства $W_B(G)$:

1. $G \subseteq W_B(G)$.
2. Если $G' \subseteq G$, то $W_B(G') \subseteq W_B(G)$.

Однотипные[1] алгебры $\langle A_1, \{f_1, f_2, f_3, \ldots\}\rangle$ и $\langle A_2, \{g_1, g_2, g_3, \ldots\}\rangle$ *гомоморфны*, если существует такое отображение $\varphi : A_1 \to A_2$ (*гомоморфизм*), что для любого i и любых $x_1, \ldots, x_n \in A_1$ выполнено $\varphi(f_i(x_1, \ldots, x_n)) = g_i(\varphi(x_1), \ldots, \varphi(x_n))$. Если отображение φ является биекцией, то оно называется *изоморфизмом*, а соответствующие алгебры — *изоморфными*.

Пусть задан гомоморфизм $\varphi : E_k \to E_r$. Для распределения $\mathbf{p}$ на E_k положим

$$\varphi(\mathbf{p}) = \left(\sum_{\varphi(i)=0} p_i, \sum_{\varphi(i)=1} p_i, \ldots, \sum_{\varphi(i)=r-1} p_i\right) \in T^{(r)}.$$

Если при этом φ — изоморфизм, то в каждой сумме в равенстве выше будет ровно одно слагаемое. Легко проверяются следующие леммы.

Лемма о гомоморфизме. *Пусть φ — гомоморфизм алгебры $\langle E_k, \Omega\rangle$ в алгебру $\langle E_r, \Omega'\rangle$. Тогда $\{\varphi(\mathbf{q}) \mid \mathbf{q} \in W_\Omega(\mathbf{p})\} = W_{\Omega'}(\varphi(\mathbf{p}))$.*

Лемма об изоморфизме. *Пусть $\langle A, \Omega\rangle$ — подалгебра алгебры $\langle E_k, \Omega\rangle$, изоморфная $\langle E_r, \Omega'\rangle$, и φ — соответствующий изоморфизм. Пусть $N(\mathbf{p}) \subseteq A$. Тогда $W_\Omega(\mathbf{p})$ в точности состоит из таких $\mathbf{q}$, что $N(\mathbf{q}) \subseteq A$, $\varphi(\mathbf{q}) \in W_{\Omega'}(\varphi(\mathbf{p}))$ и компоненты распределения $\mathbf{q}$ с индексами из множества A являются перестановкой компонент некоторого распределения из $W_{\Omega'}(\varphi(\mathbf{p}))$.*

Квазигруппы и LQ-алгебры

Алгебра $\langle E_k; \circ\rangle$ называется *квазигруппой*, если $\circ$ — бинарная операция, и для любых $a, b \in E_k$ каждое из уравнений $a \circ x = b$, $x \circ a = b$ имеет, и при том единственное, решение в E_k. Если существует такой элемент $e \in E_k$, называемый *единичным*, что $e \circ a = a \circ e = a$ для любого $a \in E_k$, то квазигруппа называется *лупой*.

Теорема 1 [2]. *Пусть $\mathbf{p} \in T^{(k)}$ — такое распределение, что $|N(\mathbf{p})| > k/2$ и пусть $\langle E_k, f\rangle$ — квазигруппа. Тогда $\left(\frac{1}{k}, \ldots, \frac{1}{k}\right) \in W_{\{f\}}(\mathbf{p})$.*

Алгебру $\langle E_k; +, \times\rangle$ будем называть *LQ-алгеброй*, если $\langle E_k; +\rangle$ — лупа, единичным элементом которой является $0 \in E_k$, $\langle E_k \setminus \{0\}; \times\rangle$ — квазигруппа, и для любого $a \in E_k$ выполнено $a \times 0 = 0 \times a = 0$.

Теорема 2 [3]. *Пусть $\mathbf{p} \in T^{(k)}$ — такое распределение, что $|N(\mathbf{p})| = k$, $\mathbf{q} \in T^{(k)}$ — произвольное распределение, и пусть $\langle E_k; f, g\rangle$ — LQ-алгебра. Тогда $\{t\mathbf{q} + (1-t)\left(\frac{1}{k}, \ldots, \frac{1}{k}\right) \mid t \in [0; 1]\} \in W_{\{f,g\}}(\{\mathbf{p}, \mathbf{q}\})$.*

[1] То есть такие, у которых совпадают количество и арность функциональных символов.

Конечные аппроксимирующие системы

Теорема 3. *Пусть $B \subset P_k$ — такое конечное множество функций, что в нем содержатся все функции одной переменной, принимающие не более двух значений, и для любого $A \subseteq E_k$ найдутся такие функции $f, g \in B$, что $\langle A; f, g\rangle$ изоморфна LQ-алгебре с основным множеством размера $|A|$. Пусть распределение $\mathbf{p} \in T^{(k)}$ таково, что $|N(\mathbf{p})| > 1$. Тогда $W_B(\mathbf{p}) = T^{(k)}$.*

Доказательство. Пусть $\mathbf{p}$ — распределение, удовлетворяющее условиям теоремы. Для каждого подмножества $A \subseteq E_k$ покажем, что любое распределение $\mathbf{q}$, у которого $N(\mathbf{q}) = A$, лежит в $W_B(\mathbf{p})$.

Для подмножеств из одного элемента это верно, так как соответствующие распределения $\mathbf{q}$ получаются подстановкой $\mathbf{p}$ в функции, индуцированные функциями из B, принимающими ровно одно значение, а по условию теоремы в множестве B лежат все такие функции.

Пусть $|A| = 2$. Поскольку $|N(\mathbf{p})| > 1$, найдутся $i, j \in N(\mathbf{p})$, $i \neq j$. Пусть $A = \{i', j'\}$. Тогда найдется такая функция $h \in B$, что $h(i) = i'$, $h(j) = j'$ (принимающая на элементах $E_k \setminus \{i, j\}$ значения i', j' произвольным образом). Положим $\mathbf{p}' = \hat{h}(\mathbf{p})$. Очевидно, что тогда $N(\mathbf{p}') = A$. Пусть $\mathbf{q}^{(1)}$ и $\mathbf{q}^{(2)}$ таковы, что $N(\mathbf{q}^{(1)}) = \{i'\}$, $N(\mathbf{q}^{(2)}) = \{j'\}$. По ранее доказанному, $\mathbf{q}^{(1)}, \mathbf{q}^{(2)} \in W_B(\mathbf{p})$.

По условию теоремы, найдутся такие функции $f, g \in B$, что $\langle A; f, g\rangle$ изоморфно LQ-алгебре на множестве E_2. В силу теоремы 2 и леммы об изоморфизме, с учетом того, что $|N(\mathbf{p}')| = 2$, получаем, что

$$W_{\{f,g\}}(\{\mathbf{p}', \mathbf{q}^{(\tau)}\}) \supseteq \{t\mathbf{q}^{(\tau)} + (1-t)(\ldots, 0, \underbrace{\frac{1}{2}}_{i'\text{-е место}}, 0, \ldots, 0, \underbrace{\frac{1}{2}}_{j'\text{-е место}}, 0, \ldots)\}, \quad \tau = 1, 2,$$

а следовательно, $W_B(\mathbf{p})$ содержит все распределения, носитель которых равен A.

Далее будем доказывать утверждение индукцией по размеру множества A. Пусть для всех A, $|A| < m$ утверждение верно, покажем, что оно также верно и для произвольного A, $|A| = m$.

Рассмотрим $A \subseteq E_k$, $|A| = m$. По условию теоремы существуют такие $f, g \in B$, что $\langle A; f, g\rangle$ изоморфно LQ-алгебре $\langle E_m; +, \times\rangle$. Пусть φ — соответствующий изоморфизм, являющийся биекцией A на E_m.

По предположению индукции, любое распределение $\mathbf{q}$, у которого $|N(\mathbf{q})| \leqslant m-1$, лежит в $W_B(\mathbf{p})$. Пусть $G = \{\varphi(\mathbf{q}) \mid \mathbf{q} \in A, |N(\mathbf{q})| \leqslant m-1\} \subset$ $\subset T^{(m)}$.

Заметим, что в силу предположения индукции, множество G содержит *все* распределения из $T^{(m)}$, носитель которых содержит менее m

элементов. Покажем, что $W_{\{+,\times\}}(G) = T^{(m)}$, тогда в силу леммы об изоморфизме будет доказано, что $W_B(\mathbf{p}) \supseteq \{\mathbf{q} \mid N(\mathbf{q}) = A\}$.

Поскольку $m > 2$, выполнено $m - 1 > m/2$, и, применяя теорему 1, получаем, что

$$W_{\{+,\times\}}(G) \ni \left(\frac{1}{m}, \dots, \frac{1}{m}\right) = \mathbf{u}.$$

Пусть далее $\mathbf{r} \in T^{(m)}$ — произвольный вектор, у которого $|N(\mathbf{r})| = m$, $\mathbf{r} \neq \mathbf{u}$. Покажем, что $\mathbf{r} \in W_{\{+,\times\}}(G)$.

Рассмотрим распределения вида $(1-t)\mathbf{u} + t\mathbf{r}$. При любом значении t такой вектор имеет сумму компонент, равную единице. При $t \in [0;1]$ все компоненты такого распределения строго положительны (в силу предположения $|N(\mathbf{r})| = m$), и поэтому вектор является стохастическим. Пусть

$$t_0 = \min\{t \mid t > 1, (1-t)\frac{1}{m} + tr_i = 0 \text{ для некоторого } i \in 0, \dots, m-1\}$$

Тогда вектор $\mathbf{q} = (1-t_0)\mathbf{u}+t_0\mathbf{r}$ имеет только неотрицательные компоненты (следовательно, является стохастическим), и при этом $|N(\mathbf{q})| < m$. Отсюда по свойству множества G вытекает, что $\mathbf{q} \in G$. По теореме 2, имеет место $\mathbf{r} \in W_{\{+,\times\}}(\{\mathbf{u}, \mathbf{q}\})$. При этом, как показано ранее, $\mathbf{u} \in W_{\{+,\times\}}(G)$ и $\mathbf{q} \in G$, поэтому $\mathbf{r} \in W_{\{+,\times\}}(G)$. Следовательно, $W_{\{+,\times\}}(G) = T^{(m)}$, откуда $\{\mathbf{q} \mid N(\mathbf{q}) = A\} \subseteq W_B(\mathbf{p})$. Доказательство шага индукции завершает доказательство теоремы. $\square$

Следствие. *Для любого $k \geqslant 2$ существует такое конечное множество $B \subset P_k$, что для любого распределения $\mathbf{p} \in T^{(k)}$ с $|N(\mathbf{p})| > 1$ имеет место $W_B(\mathbf{p}) = T^{(k)}$.*

Приведем пример конечной системы функций в P_3, позволяющей аппроксимировать любое распределение, с начальным распределением, имеющим более одной ненулевой компоненты. Система функций $B \subset P_3$ должна содержать следующие функции одной переменной:

x	$h_{0,a}$	$h_{1,a,b}$	$h_{2,a,b}$	$h_{3,a,b}$
0	a	a	a	b
1	a	a	b	a
2	a	b	a	a

для всех $a, b \in \{0, 1, 2\}$. Кроме того, в B должны лежать две функции от двух переменных, выражающие $x + y \bmod 3$ и $xy \bmod 3$, а также следующие частичные функции:

f_{01}	0	1	2
0	0	1	$*$
1	1	0	$*$
2	$*$	$*$	$*$

f_{02}	0	1	2
0	0	$*$	2
1	$*$	$*$	$*$
2	2	$*$	0

f_{12}	0	1	2
0	$*$	$*$	$*$
1	$*$	1	2
2	$*$	2	1

g_{01}	0	1	2
0	0	0	$*$
1	0	1	$*$
2	$*$	$*$	$*$

g_{02}	0	1	2
0	0	$*$	0
1	$*$	$*$	$*$
2	0	$*$	2

g_{12}	0	1	2
0	$*$	$*$	$*$
1	$*$	1	1
2	$*$	1	2

Несложно заметить, что некоторые из этих частичных функций могут быть доопредлены до одной и той же функции. Например, пара функций g_{01} и f_{12} имеет общее доопределение, а каждая из трех функций g_{01}, g_{02}, g_{12} доопределяется до $\min(x, y)$. Таким образом, количество функций в множестве B можно даже уменьшить.

Список литературы

[1] Мальцев А. И. Алгебраические системы. М.: Наука, 1970.

[2] Яшунский А. Д. О преобразованиях распределений вероятностей бесповторными квазигрупповыми формулами // Дискретная математика. 2013. Т. 25, № 2. С. 149–159.

[3] Яшунский А. Д. О бесповторных преобразованиях случайных величин над конечными полями // Дискретная математика. 2015. Т. 27, № 3. С. 145–157.