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

Препринты ИПМ • Препринт № 58 за 2016 г.

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

Яшунский А.Д.

Приближения вероятностных распределений функциями k-значной логики

Рекомендуемая форма библиографической ссылки: Яшунский А.Д. Приближения вероятностных распределений функциями k-значной логики // Препринты ИПМ им. М.В.Келдыша. 2016. № 58. 7 с. doi:10.20948/prepr-2016-58
URL: http://library.keldysh.ru/preprint.asp?id=2016-58

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

Приближения вероятностных распределений функциями k-значной логики

Москва — 2016

Яшунский А. Д.

Приближения вероятностных распределений функциями k-значной логики

Рассматриваются преобразования распределений случайных величин над конечным множеством. Показано, что, подставляя вместо переменных соответствующей функции k-значной логики независимые одинаково распределенные случайные величины, имеющие заданное распределение, можно сколь угодно точно приблизить любое наперед заданное распределение.

Ключевые слова: случайная величина, функция k-значной логики, аппроксимация

Alexey Dmitrievich Yashunsky

Approximation of probability distributions by k-valued logic functions

We consider distribution transformation for random variables over a finite set. We show that substituting independent identically distributed random variables with a given distribution for variables of a specific k-valued functions, one my obtain an arbitrarily close approximation of any given distribution.

Key words: random variable, k-valued logic function, approximation

Исследование выполнено за счет гранта Российского научного фонда (проект №14-21-00025).

Оглавление

Введение

При исследованиях многомерных дискретных операторов Шрёдингера в некоторых работах рассматривался вероятностный подход [1], в рамках которого рассматривались преобразования независимых случайных величин в результате действия дискретных операторов.

Отметим, что в случае, например, численного моделирования таких преобразований с использованием вычислительной техники неизбежно происходит дискретизация значений случайных величин (помимо дискретности пространства действия операторов, присутствующей в подобных задачах изначально). В связи с этим оправдано рассмотрение преобразований, осуществляющих отображение одних дискретных случайных величин в другие.

В рамках данной работы показывается, что дискретизация, в определенном смысле, не создает дополнительных ограничений для моделирования, а именно — подходящим образом подобранное преобразование позволяет отобразить набор независимых одинаково распределенных случайных величин в случайную величину, распределение которой сколь угодно близко к любому наперед заданному распределению.

Постановка задачи

Рассматриваются случайные величины, принимающие значения в конечном множестве $\{0, 1, \ldots, k-1\}$, которое далее обозначается E_k.

Вектор $\boldsymbol{p} = (p_0, \ldots, p_{k-1})$ будем называть *стохастическим* (или *вектором распределения*), если выполнено $p_i \geqslant 0$ при всех $i = 0, \ldots, k-1$ и $p_0 + \ldots + p_{k-1} = 1$. Каждый такой вектор задает распределение случайной величины, принимающей значения из множества $\{0, 1, \ldots, k-1\}$.

Пусть $f(x_1, \ldots, x_n)$ — функция k-значной логики. Предположим, что вместо переменных функции f подставляются независимые в совокупности случайные величины $X_1, \ldots, X_n$, имеющие распределения

$$\boldsymbol{p}^{(1)} = (p_0^{(1)}, \ldots, p_{k-1}^{(1)}), \ldots, \boldsymbol{p}^{(n)} = (p_0^{(n)}, \ldots, p_{k-1}^{(n)})$$

соответственно. Тогда вероятность того, что $f(X_1, \ldots, X_n)$ принимает значение $i \in \{0, 1, \ldots, k-1\}$, выражается следующим образом:

$$\mathsf{P}\{f(X_1, \ldots, X_n) = i\} = \sum_{(j_1, \ldots, j_n): f(j_1, \ldots, j_n) = i} p_{j_1}^{(1)} p_{j_2}^{(2)} \cdots p_{j_n}^{(n)}.$$

Можно считать, что функции k-значной логики f поставлено в соответствие вектор-отображение $\hat{f}(\boldsymbol{p}^{(1)}, \ldots, \boldsymbol{p}^{(n)})$, преобразующее набор из n стохастических векторов $\boldsymbol{p}^{(1)}, \ldots, \boldsymbol{p}^{(n)}$ в стохастический вектор

$$(\hat{f}_0(\boldsymbol{p}^{(1)}, \ldots, \boldsymbol{p}^{(n)}), \ldots, \hat{f}_{k-1}(\boldsymbol{p}^{(1)}, \ldots, \boldsymbol{p}^{(n)})).$$

Будем говорить, что стохастический вектор $\boldsymbol{u}$ *приближает* стохастический вектор $\boldsymbol{v}$ с точностью $\varepsilon > 0$, если выполнено $\max\limits_i |u_i - v_i| < \varepsilon$.

Рассматриваемая задача аппроксимации распределений заключается в том, чтобы для произвольных стохастических векторов $\boldsymbol{p}$ и $\boldsymbol{\eta}$, а также заданного произвольного $\varepsilon > 0$ построить такую функцию f, что $\hat{f}(\boldsymbol{p}, \ldots, \boldsymbol{p})$ приближает $\boldsymbol{\eta}$ с точностью ε, т. е.

$$\max_i |\hat{f}_i(\boldsymbol{p}, \ldots, \boldsymbol{p}) - \eta_i| < \varepsilon.$$

Отметим, что если какая-то из компонент вектора $\boldsymbol{p}$ равна 1, то все прочие компоненты равны 0; такие векторы будем называть *вырожденными*. Каково бы ни было отображение f, если вектор $\boldsymbol{p}$ вырожденный, то и вектор $\hat{f}(\boldsymbol{p}, \ldots, \boldsymbol{p})$ также вырожденный. Отсюда легко видеть, что задачу аппроксимации распределений имеет смысл рассматривать только для невырожденных векторов $\boldsymbol{p}$.

В данной работе показывается, что задача аппроксимации имеет решение в классе всевозможных функций k-значной логики.

Вспомогательные утверждения

Для стохастического вектора $\boldsymbol{p} = (p_0, \ldots, p_{k-1})$ будем называть множество $N(\boldsymbol{p}) = \{i \in E_k \mid p_i \neq 0\}$ *носителем* распределения $\boldsymbol{p}$.

Покажем, что с помощью функций k-значной логики можно преобразовывать распределения, расширяя их носитель.

Лемма 1. *Пусть $\boldsymbol{p}$ — такой стохастический вектор, что $|N(\boldsymbol{p})| > 1$. Тогда существует такая k-значная функция h, что $N(\hat{h}(\boldsymbol{p}, \ldots, \boldsymbol{p})) = E_k$.*

Доказательство. Пусть $|N(\boldsymbol{p})| = m$, $1 < m < k$. Покажем, что существует такая функция χ, что $|N(\hat{\chi}(\boldsymbol{p}, \ldots, \boldsymbol{p}))| = m + 1$. Тогда индукцией по m получим, что существует такая функция h, что $|N(\hat{h}(\boldsymbol{p}, \ldots, \boldsymbol{p}))| = k$, что равносильно утверждению леммы.

Определим функцию χ. Пусть $i \in E_k \setminus N(\boldsymbol{p})$. Поскольку $m < k$, такое значение i существует. Положим

$$\chi(x_1, x_2) = \begin{cases} x_1, \text{ если } x_1 = x_2 \\ i, \text{ иначе.} \end{cases}$$

Несложно проверить, что такая функция обладает требуемым свойством. $\square$

Лемма 2. *Пусть $\boldsymbol{p}$ — такой стохастический вектор, что $N(\boldsymbol{p}) = E_k$. Тогда для любого $\varepsilon > 0$ существует такая k-значная функция g, что*

$$\max_i \left| \hat{g}_i(\boldsymbol{p}, \ldots, \boldsymbol{p}) - \frac{1}{k} \right| < \varepsilon.$$

Доказательство. Рассмотрим функции

$$g^{(n)}(x_1, \ldots, x_n) = x_1 + \ldots + x_n (\operatorname{mod} k).$$

Положим $\boldsymbol{\gamma}^{(n)} = \hat{g}^{(n)}(\boldsymbol{p}, \ldots, \boldsymbol{p})$. Несложно проверить, что

$$(\gamma_0^{(n)}, \ldots, \gamma_{k-1}^{(n)}) =$$
$$= (\gamma_0^{(n-1)}, \ldots, \gamma_{k-1}^{(n-1)}) \begin{pmatrix} p_{0-0 \bmod k} & p_{1-0 \bmod k} & \cdots & p_{k-1-0 \bmod k} \\ p_{0-1 \bmod k} & p_{1-1 \bmod k} & \cdots & p_{k-1-1 \bmod k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{0-(k-1) \bmod k} & p_{1-(k-1) \bmod k} & \cdots & p_{k-1-(k-1) \bmod k} \end{pmatrix},$$

причем $\boldsymbol{\gamma}^{(1)} = \boldsymbol{p}$. Таким образом распределения $\boldsymbol{\gamma}^{(n)}$ соответствуют состояниям цепи Маркова с матрицей $\{p_{(i-j) \bmod k}\}_{ij}$. Несложно проверить, что при $N(\boldsymbol{p}) = E_k$ такая цепь является эргодической, а ее предельное распределение равномерно (в силу того, что матрица $\{p_{i-j}\}_{ij}$ бистохастическая, см., например, [2]).

Итак, $\boldsymbol{\gamma}^{(n)} \to \left(\frac{1}{k}, \ldots, \frac{1}{k}\right)$ при $n \to \infty$, а следовательно, для любого $\varepsilon > 0$ найдется такое m, что

$$\max_i \left| \hat{g}_i^{(m)}(\boldsymbol{p}, \ldots, \boldsymbol{p}) - \frac{1}{k} \right| < \varepsilon.$$

$\square$

Отметим, что используемое в доказательстве леммы 2 свойство функций $x_1 + \ldots + x_n (\operatorname{mod} k)$ относится к математическому фольклору. Более сильная версия этого свойства, относящаяся к произвольным суперпозициям квазигрупповых операций, доказана в работе [3].

Теорема приближения

Теорема. *Пусть $\boldsymbol{p} = (p_0, p_1, \ldots, p_{k-1})$ — k-мерный стохастический вектор, причем $p_i \neq 1$ для любого i, и пусть $\boldsymbol{\eta} = (\eta_0, \eta_1, \ldots, \eta_{k-1})$ — произвольный k-мерный стохастический вектор. Тогда для любого $\varepsilon > 0$ существует* *такая функция $f \in P_k$, что*

$$\max_i |\hat{f}_i(\boldsymbol{p}, \ldots, \boldsymbol{p}) - \eta_i| < \varepsilon.$$

Доказательство. Пусть $\boldsymbol{\eta}$ — стохастический вектор, который требуется аппроксимировать с точностью $\varepsilon > 0$. Рассмотрим φ — функцию k-значной логики от n переменных. Определим значения φ следующим образом: на $\lfloor \eta_0 \cdot k^n \rfloor$ наборах положим ее равной 0, на $\lfloor \eta_1 \cdot k^n \rfloor$ положим ее равной 1, и так далее: на $\lfloor \eta_i \cdot k^n \rfloor$ наборах положим функцию равной i для всех $i = 0, \ldots, k-2$. На всех оставшихся наборах положим функцию φ равной $k-1$.

Пусть $\boldsymbol{u} = \left(\frac{1}{k}, \ldots, \frac{1}{k}\right)$ — стохастический вектор равномерного распределения. Рассмотрим при $i \leqslant k-2$ величину

$$|\hat{\varphi}_i(\boldsymbol{u}, \ldots, \boldsymbol{u}) - \eta_i| = \left| \frac{\lfloor \eta_i \cdot k^n \rfloor}{k^n} - \eta_i \right| \leqslant \frac{1}{k^n}.$$

Для $i = k-1$ имеем

$$|\hat{\varphi}_{k-1}(\boldsymbol{u}, \ldots, \boldsymbol{u}) - \eta_{k-1}| = \left| 1 - \sum_{i=0}^{k-2} \hat{\varphi}_i(\boldsymbol{u}, \ldots, \boldsymbol{u}) - 1 + \sum_{i=0}^{k-2} \eta_i \right| \leqslant$$
$$\leqslant \sum_{i=0}^{k-2} |\hat{\varphi}_i(\boldsymbol{u}, \ldots, \boldsymbol{u}) - \eta_i| \leqslant \frac{k-1}{k^n} \leqslant \frac{1}{k^{n-1}}.$$

Отсюда $\max\limits_i |\hat{\varphi}_i(\boldsymbol{u}, \ldots, \boldsymbol{u}) - \eta_i| \leqslant \frac{1}{k^{n-1}}$. Пусть далее n таково, что $\frac{1}{k^{n-1}} < \frac{\varepsilon}{2}$.

Пусть $\boldsymbol{p}$ — стохастический вектор, удовлетворяющий условиям теоремы. По лемме 1 существует такая функция h, что $N(\hat{h}(\boldsymbol{p}, \ldots, \boldsymbol{p})) = E_k$. Положим $\boldsymbol{r} = \hat{h}(\boldsymbol{p}, \ldots, \boldsymbol{p})$.

Из непрерывности $\hat{\varphi}$ по каждой из своих переменных легко вывести, что для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для любого стохастического вектора $\boldsymbol{v}$, удовлетворяющего $\max\limits_i |v_i - u_i| < \delta$, имеет место

$$\max_i |\hat{\varphi}_i(\boldsymbol{v}, \ldots, \boldsymbol{v}) - \hat{\varphi}_i(\boldsymbol{u}, \ldots, \boldsymbol{u})| < \frac{\varepsilon}{2}.$$

В силу леммы 2 найдется такая функция g, что $\max\limits_i |\hat{g}(\boldsymbol{r}, \ldots, \boldsymbol{r})_i - u_i| \leqslant$ $\leqslant \delta$. Положим $\boldsymbol{v} = \hat{g}(\boldsymbol{r}, \ldots, \boldsymbol{r})$. Тогда в силу выбора функции g имеет место $\max\limits_i |\hat{\varphi}_i(\boldsymbol{v}, \ldots, \boldsymbol{v}) - \hat{\varphi}_i(\boldsymbol{u}, \ldots, \boldsymbol{u})| < \frac{\varepsilon}{2}$. Объединяя полученные выше неравенства, имеем

$$\max_i |\hat{\varphi}_i(\boldsymbol{v}, \ldots, \boldsymbol{v}) - \eta_i| \leqslant$$
$$\leqslant \max_i |\hat{\varphi}_i(\boldsymbol{v}, \ldots, \boldsymbol{v}) - \hat{\varphi}_i(\boldsymbol{u}, \ldots, \boldsymbol{u})| + \max_i |\hat{\varphi}_i(\boldsymbol{u}, \ldots, \boldsymbol{u}) - \eta_i| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Отсюда следует, что искомая функция f равна

$$\varphi(g(h(x_1,\ldots),\ldots,h(\ldots)),\ldots,g(h(\ldots),\ldots,h(\ldots,x_M))),$$

где все переменные, подставляемые в формулу, различны. $\square$

Список литературы

[1] Faustino N. Hypercomplex Fock States for Discrete Electromagnetic Schrödinger Operators: A Bayesian Probability Perspective // arXiv:1505.05926 [math-ph].

[2] Saloff-Coste L. Random walks on finite groups // Probability on discrete structures. Encyclopaedia Math. Sci., 110. Springer, Berlin, 2004. Pp. 263–346.

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