Чувствительность оценки вероятности свидетельства локального апостериорного вывода в алгебраических байесовских сетях: вычислительные эксперименты

Е. А. Мальчевская[1], В. Ф. Столярова[2]
СПИИРАН
[1]katerina.malch@gmail.com, [2]valerie.stoliarova@gmail.com

Аннотация. **В статье представлены результаты вычислительных экспериментов по анализу чувствительности первого уравнения локального апостериорного вывода для набора детерминированных свидетельств в алгебраических байесовских сетях. Для полученных оценок приведены графики зависимости колебания результата от исходной вариации оценок вектора вероятностей скалярного фрагмента знаний. На языке CSharp в программном комплексе AlgBN Math Library для логико-вероятностного вывода был реализован модуль для вычисления оценок. Также результаты экспериментов были визуализированы с помощью таблицы. Предложенная визуализация доказывает высокую чувствительность используемых моделей, что подтверждает корректность их использования.**

Ключевые слова: вероятностные графические модели; алгебраические байесовские сети; апостериорный вывод; детерминированное свидетельство

I. Введение

Алгебраические байесовские сети (АБС) являются одним из классов вероятностных графических моделей и позволяют обрабатывать данные с неопределенностью. Данные могут содержать неопределенность сами по себе, как например, информация, получаемая на естественном языке, или же содержать пропуски. Таким образом, при работе с сетью, могут поступать стохастические, детерминированные и неточные (интервальные) свидетельства. АБС используют аппарат логико-вероятностного вывода [1, 4]. Возможность работать с неопределенностью различного рода является одним из преимуществ алгебраических байесовских сетей (АБС) в сравнении с другими родственными им вероятностными графическими моделями [6]. Логико-вероятностный вывод используется для решения задач проверки и поддержания непротиворечивости, априорного вывода и апостериорного вывода. Формализация различных задач АБС была произведена [4, 5].

Работа выполнена при финансовой поддержке РФФИ, проект №18-01-00626 и по госбюджетной теме № 0073-2018-0001

В рамках данной статьи обратимся к первой задаче апостериорного вывода, матрично-векторные уравнения для которой были приведены в [3]. Первая задача апостериорного вывода сводится к вычислению вероятности свидетельства на основе оценок вероятности истинности, заданных во фрагменте знаний. Эта задача может рассматриваться как на уровне всей сети, так и на локальном уровне, в рамках данной работы рассмотрен локальный вывод.

Важным элементом работы с моделью АБС является оценка чувствительности модели к различным колебаниям исходных данных [3]. Анализ чувствительности является одним из критериев оценки математической модели [8, 9]. Данная оценка характеризует степень изменения результата в зависимости от колебания значений входных данных.

Пусть имеется фрагмент знаний АБС, в который поступает набор детерминированных свидетельств. Целью статьи является оценка чувствительности уравнения для решения первой задачи апостериорного вывода к вариации вероятностей элементов идеала конъюнктов фрагмента знаний. Исследование осуществляется при помощи вычислительных экспериментов на языке C# в программном комплексе AlgBN Math Library [2].

II. Основные понятия

Определения и понятия далее излагаются в соответствии с [6]. Алгебраические байесовские сети (АБС) представляются ненаправленными графами с идеалами конъюнктов в узлах. Конъюнктам приписывается либо скалярная (точечная), либо интервальная оценка вероятности истинности. Назовем идеал конъюнктов с приписанными его элементам оценками вероятности истинности фрагментом знаний (ФЗ).

Как и другие пропозициональные формулы, идеал конъюнктов строится над некоторым фиксированным множеством атомов – алфавитом. Оценки вероятности образуют вектор P_c.

Под детерминированным свидетельством будем понимать предположение о том, что один или несколько атомов получили конкретные означивания. Детерминированное свидетельство представляется в виде пары $\langle c_i, c_j \rangle$, где c_i – квант, содержащий все положительно означенные атомы свидетельства, c_j – квант, содержащий все отрицательно означенные атомы свидетельства.

Вычисление вероятности поступающего свидетельства в рамках первой задачи локального апостериорного вывода в АБС для детерминированного свидетельства $\langle c_i, c_j \rangle$ осуществляется по формуле [7]:

$$p\left(\langle c_i, c_j \rangle\right) = \left(\mathrm{r}^{\langle i,j \rangle}, \mathrm{P_c}\right),$$

где

$$\mathrm{r}^{\langle i,j \rangle} = \tilde{\mathrm{r}}_{n-1}^{\langle i,j \rangle} \otimes \tilde{\mathrm{r}}_{n-2}^{\langle i,j \rangle} \otimes \cdots \tilde{\mathrm{r}}_{0}^{\langle i,j \rangle},$$

$$\tilde{\mathrm{r}}_{k}^{\langle i,j \rangle} = \begin{cases} \mathrm{r}^{+}, x_k \in c_i \\ \mathrm{r}^{-}, x_k \in c_i\,, \\ \mathrm{r}^{\mathrm{o}}, \textit{иначе} \end{cases}$$

$$\mathrm{r}^{+} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \mathrm{r}^{-} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}; \mathrm{r}^{\mathrm{o}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

III. Задача линейного программирования для оценки чувствительности и вычислительные эксперименты

A. *Задача линйного программирования*

Итак, пусть во фрагмент знаний поступает детерминированное свидетельство $\langle c_i, c_j \rangle$. Пусть исходные оценки истинности фрагмента знаний $\mathrm{P_c}$ могут варьироваться не больше, чем на $\delta : v\left(\mathrm{P_c^o}, \hat{\mathrm{P}}_\mathrm{c}\right) \le \delta$. Тогда чувствительность $\varepsilon(\mathrm{P_c^o})$ первой задачи апостериорного вывода [10] может быть получена в результате решения задачи линейного программирования:

$$\varepsilon(\mathrm{P_c^o}) = \max_{\substack{\hat{\mathrm{P}}_\mathrm{c}\mathrm{I}_n \ge 0, \mathrm{P_c^o}\mathrm{I}_n \ge 0, \\ v\left(\mathrm{P_c^o}, \hat{\mathrm{P}}_\mathrm{c}\right) \le \delta, \\ p\left(\langle c_i, c_j \rangle\right) = \left(\mathrm{r}^{\langle i,j \rangle}, \mathrm{P_c^o}\right), \\ \hat{p}\left(\langle c_i, c_j \rangle\right) = \left(\mathrm{r}^{\langle i,j \rangle}, \hat{\mathrm{P}}_\mathrm{c}\right)}} \left\{ p - \hat{p}, \hat{p} - p \right\}. \tag{1}$$

В первом блоке ограничений задаются условия на непротиворечивость оценок вероятности истинности элементов рассматриваемого вектора P_c. С помощью второго блока фиксируются скалярные оценки исходного вектора. В третьем блоке задаются условия на непротиворечивость варьируемого вектора $\hat{\mathrm{P}}_c$. Четвертый – накладывает ограничения на вариативность. Следующие два ограничения задают то, каким образом решения первой задачи апостериорного вывода для фиксированного свидетельства и векторов P_c и $\hat{\mathrm{P}}_c$ выражаются через элементы этих векторов соответственно. Последнее условие гарантирует нам выполнение необходимого условия того, что вероятность пустой конъюнкции в варьируемом векторе будет равна 1.

Для исследования поведения оценки чувствительности первой задачи апостериорного вывода для детерминированного свидетельства при колебании входных оценок вероятности истинности элементов исходного ФЗ были проведены вычислительные эксперименты.

B. *Вычислительные экперименты*

Зададим ФЗ $\left(C, \mathrm{P}_{c_0}\right)$ над алфавитом из двух атомов $A = \{x_2, x_1\}$. Следующий вектор состоит из скалярных оценок ФЗ:

$$\mathrm{P}_c^{\mathrm{o}} = \begin{pmatrix} 1 \\ p(x_1) \\ p(x_2) \\ p(x_2 x_1) \end{pmatrix} = \begin{pmatrix} 1 \\ 0.6 \\ 0.5 \\ 0.333 \end{pmatrix}$$

Для заданного ФЗ будем решать первую задачу апостериорного вывода со следующими детерминированными свидетельствами:

$$\langle x_1 \rangle; \langle , x_1 \rangle; \langle x_2 \rangle; \langle x_2, x_1 \rangle; \langle x_1, x_2 \rangle.$$

ЗАДАЧА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ ДЛЯ ОЦЕНКИ ЧУ

Перечисленные выше свидетельства соответствуют следующим пропозициональным формулам:

$$x_1; \bar{x}_1; x_2; x_2\bar{x}_1; \bar{x}_2 x_1.$$

Данным свидетельствам соответствуют следующие вектора-редистрибьютеры:

$$\mathrm{r}^{\langle x_1 \rangle} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}; \mathrm{r}^{\langle \bar{x}_1 \rangle} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}; \mathrm{r}^{\langle x_2 \rangle} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix};$$

$$\mathrm{r}^{\langle x_2, x_1 \rangle} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}; \mathrm{r}^{\langle x_1, x_2 \rangle} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}.$$

Для каждого свидетельства и фиксированного δ вычислим решим задачу линейного программирования (ЗЛП) (1).

ТАБЛИЦА I ОЦЕНКИ ЧУВСТВИТЕЛЬНОСТИ ДЛЯ ДЕТЕРМИНИРОВАННЫХ СВИДЕТЕЛЬСТВ

Свидетельство	$\delta =$ .01	.07	.15	.27	.43	.58	.7	1
$\langle x_2, x_1 \rangle$	.02	.14	.3	.503	.663	.813	.833	.833
$\langle x_2 \rangle$	.01	.07	.15	.27	.4	.4	.4	.4
$\langle x_1, x_2 \rangle$	.02	.14	.3	.503	.663	.733	.733	.733
$\langle x_1 \rangle$	.01	.07	.15	.27	.43	.5	.5	.5
$\langle , x_1 \rangle$	.01	.07	.15	.27	.43	.5	.5	.5

В таблице приведены оценки чувствительности для фиксированных значений δ и набора рассматриваемых детерминированных свидетельств. На некоторых интервалах значений дельты прослеживается линейная зависимость дельты и получаемой оценки чувствительности.

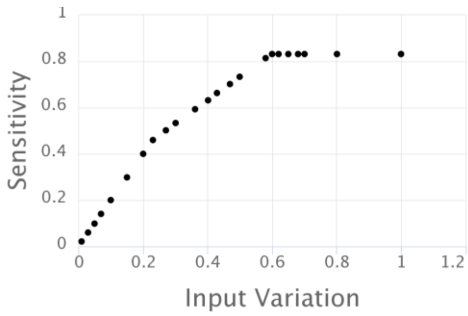

Рис. 1. График оценок чувствительности для детерминированного свидетельства $\langle x_2, x_1 \rangle$

На рис. 1 представлен график роста оценки чувствительности с ростом допустимой вариации исходного вектора в рамках рассматриваемой первой задачи апостериорного вывода для детерминированного свидетельства $\langle x_2, x_1 \rangle$ и кванта $x_2\overline{x}_1$, в виде которого она может быть представлена.

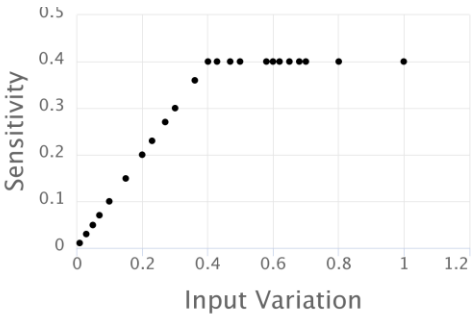

Рис. 2. График оценок чувствительности для детерминированного свидетельства $\langle x_2 \rangle$

На рис. 2 представлен график роста оценки чувствительности с ростом допустимой вариации исходного вектора в рамках рассматриваемой первой задачи апостериорного вывода для детерминированного свидетельства $\langle x_2 \rangle$ и кванта x_2, в виде которого она может быть представлена.

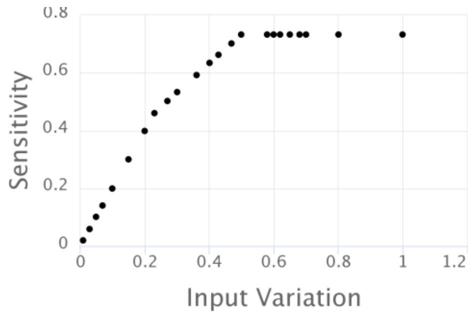

Рис. 3. График оценок чувствительности для детерминированного свидетельства $\langle x_1, x_2 \rangle$

На рис. 3 представлен график роста оценки чувствительности с ростом допустимой вариации исходного вектора в рамках рассматриваемой первой задачи апостериорного вывода для детерминированного свидетельства $\langle x_1, x_2 \rangle$ и кванта $\overline{x}_2 x_1$, в виде которого она может быть представлена.

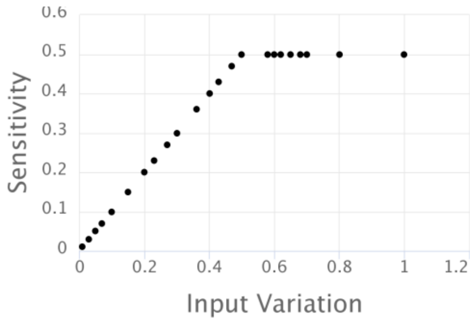

Рис. 4. График оценок чувствительности для детерминированных свидетельств $\langle x_1 \rangle$ и $\langle , x_1 \rangle$

На рис. 4 представлен график роста оценки чувствительности с ростом допустимой вариации исходного вектора в рамках рассматриваемой первой задачи апостериорного вывода для детерминированных свидетельств $\langle x_1 \rangle$ и $\langle \overline{x}_1 \rangle$, с квантами x_1 и $\overline{x}_1$ соответственно. Для обоих случаев график зависимости, как было показано экспериментально, совпал.

Как можно видеть из приведенных графиков, зависимость оценки чувствительности результата решения задачи первого вывода для представленных свидетельств выражается линейно на некоторых интервалах. Показано, что чувствительность возрастает при возрастании параметра δ, а потом стабилизируется.

IV. ЗАКЛЮЧЕНИЕ

В результате статистических экспериментов было получено, что степень чувствительности уравнения решения первой задачи апостериорного вывода…. Анализ оценок чувствительности апостериорного вывода развивает теорию АБС в частности и область машинного обучения в целом. В дальнейшем планируется, развивая полученные результаты, исследовать оценку чувствительности для стохастических и неточных свидетельств.

Список литературы

[1] Золотин А.А., Мальчевская Е.А., Харитонов Н.А., Тулупьев А.Л. Локальный и глобальный логико-вероятностный вывод в алгебраических байесовских сетях: матрично-векторное описание и вопросы чувствительности. Нечеткие системы и мягкие вычисления. 2017,12(2), 133–50.

[2] Мальчевская Е.А., Золотин А.А. Логико-вероятностный вывод в АБС: архитектура и примеры использования программного комплекса на языке C# // 3-я Всероссийская Поспеловская конференция с международным участием «Гибридные и синергетические интеллектуальные системы». Под редакцией А.В. Колесникова. 2016. С. 181–187.

[3] Сироткин А.В. Алгебраические байесовские сети: вычислительная сложность алгоритмов логико-вероятностного вывода в условиях неопределенности: Дисс.… канд. физ.-мат. наук. СПб.: СПбГУ, 2011.

[4] Сироткин А.В., Тулупьев А.Л. Матричные уравнения локального логико-вероятностного вывода в алгебраических байесовских сетях // Труды СПИИРАН. 2008. Т. 6. С. 134–143.

[5] Тулупьев А.Л. Алгебраические байесовские сети: локальный логико-вероятностный вывод: Учеб. пособие. СПб.: СПбГУ; ООО Издательство «Анатолия», 2007. 80 с. (Сер. Элементы мягких вычислений).

[6] Тулупьев А., Николенко С., Сироткин А. Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.

[7] Тулупьев А.Л., Сироткин А.В., Золотин А.А., Матричные уравнения нормирующих множителей в локальном апостериорном выводе оценок истинности в алгебраических байесовских сетях, Вестник Санкт-Петербургского университета. Серия 1. Математика. Механика. Астрономия, 2015, Т. 2(3).

[8] Dalir, F. and Motlagh, M.S. and Ashrafi, K., Sensitivity analysis of parameters affecting carbon footprint of fossil fuel power plants based on life cycle assessment scenarios, Global Journal of Environmental Science and Management, 2017, vol.3(1), pp. 75–88.

[9] Depaoli S, Yang Y, Felt J. Using bayesian statistics to model uncertainty in mixture models: a sensitivity analysis of priors, Structural Equation Modeling-a Multidisciplinary Journal, 2017, vol. 24 (2), pp. 198–215.

[10] Zolotin A. A., Malchevskaya E. A., Tulupyev A. L., Sirotkin A. V., An Approach to Sensitivity Analysis of Inference Equations in Algebraic Bayesian Networks, In International Conference on Intelligent Information Technologies for Industry, Springer, Cham, pp. 34–42.