ПРОПАГАЦИЯ ВИРТУАЛЬНОГО СВИДЕТЕЛЬСТВА В АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЯХ: АЛГОРИТМЫ И ИХ ОСОБЕННОСТИ[1]

Золотин А. А., аспирант мат.-мех. фак-та кафедры информатики СПбГУ, стажер-исследователь Санкт-Петербургского института информатики и автоматизации РАН (СПИИРАН), andrey.zolotin@gmail.com

Шляк А. В., студент-бакалавр мат.-мех. фак-та кафедры информатики СПбГУ, стажер-исследователь Санкт-Петербургского института информатики и автоматизации РАН (СПИИРАН), dieseltsame@gmail.com

Тулупьев А.Л., д.ф.-м.н., профессор кафедры информатики СПбГУ, заведующий лабораторией ТиМПИ СПИИРАН, alt@iias.spb.su

Аннотация

В статье описан алгоритм глобального апостериорного вывода в алгебраических байесовских сетях. Подробно рассмотрены основные этапы алгоритма пропагации виртуального свидетельства между двумя фрагментами знаний над идеалами конъюнктов и формирование матрицы перехода от вектора элементов фрагмента знаний к вектору элементов виртуального свидетельства. Приведен пример использования имеющихся теоретических результатов. Поставлены задачи дальнейшего улучшения полученных результатов.

Введение

Алгебраические байесовские сети были введены профессором В. И. Городецким [1] и являются математической моделью, позволяющей обрабатывать данные с неопределенностью, которая может появляться, например, из-за различающихся оценок какого-либо утверждения экспертами. Алгебраическая байесовская сеть (АБС) состоит из набора случайных элементов, соответствующих высказываниям, которым приписаны оценки вероятностей, и структуры связи между указанными

---

[1] В работе представлены результаты исследований, поддержанные грантом РФФИ 15-01- 09001, рук. А.Л. Тулупьев

высказываниями, которую обычно представляют в виде графа смежности. Родственный АБС класс байесовских сетей доверия находит применение во множестве областей науки и промышленности, таких как оценка рисков [9, 10, 12, 13] или системы поддержки принятия решений [11, 14].

Алгоритмы пропагации виртуальных свидетельств, предложенные в [2], опираются на уравнения локального апостериорного логико-вероятностного вывода (ЛВВ)[5]. В данной работе рассматривается результат распространения матрично-векторного подхода, апробированного в локальном ЛВВ, на глобальный апостериорный вывод, а также приводится пример использования данного алгоритма. Однако рассмотренные результаты не являются чистым матрично-векторным уравнением. Целью данной статьи является постановка задач, которые требуется рассмотреть в дальнейшем для получения такого уравнения. Решение данных задач позволит исследовать чувствительность и устойчивость предложенной модели и упростит программную реализацию данных алгоритмов. В контексте статьи рассматривается связанная непротиворечивая АБС, сформированная над множеством идеалов конъюнктов, в обозначениях, принятых в [6].

Глобальный апостериорный вывод

Рассмотрим связную ациклическую АБС с фрагментами знаний, представленными идеалами конъюнктов со скалярными оценками. Пусть в один из фрагментов знаний поступило стохастическое свидетельство. Задача глобального апостериорного вывода — пропагировать это свидетельство во все фрагменты знаний сети.

Алгоритм пропагации состоит из 3 шагов:

1. Пропагация свидетельства во фрагмент знаний, в который оно пришло, и оценка апостериорных вероятностей его элементов;
2. Формирование виртуального свидетельства;
3. Пропагация виртуального свидетельства в соседний фрагмент знаний.

Аналогичным образом свидетельство пропагируется далее, пока не будут переозначены оценки во всех фрагментах знаний.

Виртуальным свидетельством называется пересечение двух фрагментов знаний (сепаратор), также являющийся фрагментом знаний. Так как после переозначивания оценок в первом фрагменте знаний, конъюнкты, принадлежащие сепаратору, имеют новые оценки, потому что принадлежат первому фрагменту знаний, но с другой стороны, они также принадлежат второму фрагменту знаний и имеют другие оценки, то сепаратор можно рассмотреть как новое свидетельство, поступившее во второй фрагмент знаний.

Таким образом, на втором шаге алгоритма из вектора вероятностей первого фрагмента знаний необходимо выделить вектор, содержащий элементы, принадлежащие обоим фрагментам знаний, и принять его за новое свидетельство, которое на третьем шаге пропагируется во второй фрагмент знаний, оценки которого необходимо переозначить.

Для того чтобы выделить необходимые значения из фрагмента знаний, построим матрицу перехода $\mathbf{Q}$ от вектора вероятностей элементов первого ФЗ — $\mathbf{P}_{\mathrm{c}}^{1}$ к вектору вероятностей элементов сепаратора — $\mathbf{P}_{\mathrm{c}}^{\mathrm{sep}}$. Известно, что $\mathbf{P}_{\mathrm{c}}^{1} = \mathbf{J}_{n}\mathbf{P}_{\mathrm{q}}^{1}$ [7]. При умножении i-той строки $\mathbf{J}_{n}[i]$ на $\mathbf{P}_{\mathrm{q}}^{1}$, получим i-тый элемент вектора $\mathbf{P}_{\mathrm{c}}^{1}$. Выделим в отдельную матрицу строки $\mathbf{J}_{n}[i]$, выделяющие из $\mathbf{P}_{\mathrm{c}}^{1}$ элементы сепаратора, чтобы, умножив её на $\mathbf{P}_{\mathrm{q}}^{1}$, получить вектор оценок вероятностей элементов виртуального свидетельства.

Для того чтобы выделить строки $\mathbf{J}_{n}[i]$, нужно домножить $\mathbf{J}_{n}$ слева на матрицу $\mathbf{Q}$ размерности $m \times n$ (m — длина вектора $\mathbf{P}_{\mathrm{c}}^{\mathrm{sep}}$, n — длина вектора $\mathbf{P}_{\mathrm{c}}^{1}$). Элементы матрицы можно определить по следующему правилу: $\mathbf{Q}[i,j] = \begin{cases} 1, \mathbf{P}_{\mathrm{c}}^{1}[j] = \mathbf{P}_{\mathrm{c}}^{sep}[i] \\ 0, \mathit{иначе} \end{cases}$. Под $\mathbf{P}_{\mathrm{c}}^{1}[i]$ и $\mathbf{P}_{\mathrm{c}}^{\mathrm{sep}}[i]$ подразумеваются сами конъюнкты, а не значения оценок вероятностей. Наконец, чтобы сформировать вектор вероятностей сепаратора, матрицу $\mathbf{Q}$ нужно умножить на вектор $\mathbf{P}_{\mathrm{q}}^{1}$: $\mathbf{P}_{\mathrm{c}}^{\mathrm{sep}} = \mathbf{Q}\mathbf{J}_{n}\mathbf{P}_{\mathrm{q}}^{1} = \mathbf{Q}\mathbf{P}_{\mathrm{c}}^{1}$.

Алфавит, над которым построен сепаратор, можно найти как $A_{\mathrm{sep}} = A_1 \cap A_2$, где A_1 — алфавит первого фрагмента знаний, A_2 — алфавит второго фрагмента знаний.

Таким образом, второй и третий шаги алгоритма описываются с помощью следующего уравнения для пропагации виртуального свидетельства и переозначивания оценок вероятностей элементов во фрагменте знаний:

$$\mathbf{P}_{\mathrm{c}}^{2,a} = \sum_{i=0}^{2^{n'}-1} \frac{\mathbf{T}^{\langle \mathrm{GInd}\,(i,m), \mathrm{GInd}\,(2^{n'}-1-i,m)\rangle}\mathbf{P}_{\mathrm{c}}^{2}}{(\mathbf{r}^{\langle \mathrm{GInd}\,(i,m), \mathrm{GInd}\,(2^{n'}-1-i,m)\rangle}, \mathbf{P}_{\mathrm{c}}^{2})} \mathbf{I}_{n}\mathbf{P}_{\mathrm{c}}^{\mathrm{sep}}[i],$$

где $\mathbf{P}_{\mathrm{c}}^{\mathrm{sep}} = \mathbf{Q}\mathbf{P}_{\mathrm{c}}^{1}$; $\mathbf{P}_{\mathrm{c}}^{1}$ и $\mathbf{P}_{\mathrm{c}}^{2}$ — векторы вероятностей элементов идеалов конъюнктов первого и второго фрагментов знаний; $\mathbf{P}_c^{2,a}$ — вектор апостерионых вероятностей элементов идеала конъюнктов второго фрагмента знаний;

$\mathrm{GInd}(i,m)$ — функция, которая по индексу наибольшего элемента C^{ev} в алфавите A и индексу конъюнкта в алфавите A^{ev} возращает индекс соответсвующего конъюнкта в алфавите A

Однако, в данной формулировке присутствует часть операций, не относящаяся к матрично-векторным вычислениям. А именно поэлементное формирование матрицы $\mathbf{Q}$ и участие в вычислении функции GInd.

Можно заметить следующий результат: обозначим за $\mathrm{c}_i = A_1 \cap A_2$ – множество тех элементов, которые входят в виртуальное свидетельство. Обозначим $\mathrm{c}_j = A_1 \setminus c_i = A_1 \setminus (A_1 \cap A_2)$ – множество тех элементов, которые входят во фрагмент знаний, но не входят в виртуальное свидетельство, A_1 и A_2 – алфавиты первого и второго ФЗ соответственно. Тогда можно представить матрицу $\mathbf{Q}$ через Кронекерово произведение: $\mathbf{Q} = \tilde{\mathbf{Q}}_{n-1}^{(i,j)} \otimes \tilde{\mathbf{Q}}_{n-2}^{(i,j)} \otimes \ldots \tilde{\mathbf{Q}}_0^{(i,j)}$, где $\tilde{\mathbf{Q}}_k^{(i,j)} = \begin{cases} \mathbf{Q}^+, \text{если } x_k \text{ входит в } c_i \\ \mathbf{Q}^-, \text{если } x_k \text{ входит в } c_j \end{cases}$

а $\mathbf{Q}^+ = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$; $\mathbf{Q}^- = (1\ 0)$, x_k – k – тый элемент в A_1.

Таким образом, актуальной задачей является доказательство того, что предложенный выше алгоритм формирования матрицы $\mathbf{Q}$ действительно дает необходимую матрицу перехода от вектора $\mathbf{P}_c^1$ к вектору $\mathbf{P}_c^{\mathrm{sep}}$.

Пример

Пусть даны два фрагмента знаний над алфавитами $A_1 = \{x_1, x_2\}$ и $A_2 = \{x_1, x_3\}$ и пусть в первый фрагмент знаний поступило стохастическое свидетельство. Переозначим оценки в данных ФЗ с учетом поступившего свидетельства. Пусть вероятности элементов ФЗ до пропагации свидетельства были следующими:

$$\mathbf{P}_c^1 = \begin{pmatrix} 1 \\ p(x_1) \\ p(x_2) \\ p(x_2x_1) \end{pmatrix} = \begin{pmatrix} 1 \\ 0.8 \\ 0.7 \\ 0.6 \end{pmatrix}; \mathbf{P}_c^2 = \begin{pmatrix} 1 \\ p(x_1) \\ p(x_3) \\ p(x_3x_1) \end{pmatrix} = \begin{pmatrix} 1 \\ 0.4 \\ 0.5 \\ 0.1 \end{pmatrix}.$$

Пусть поступившее свидетельство сформировано над алфавитом

$A_{\mathrm{ev}} = \{x_2\}$ и ему сопоставлен вектор $\mathbf{P}_{\mathrm{c}}^{\mathrm{ev}} = \begin{pmatrix} 1 \\ p(x_2) \end{pmatrix} = \begin{pmatrix} 1 \\ 0.4 \end{pmatrix}$.

Воспользуемся формулой пропагации свидетельства во фрагмент знаний над идеалом конъюкнтов [3] и вычислим значения всех в нее входящих элементов.

$m = 10,\ i = 0 : \mathrm{GInd}(0,10) = 00;\ \mathrm{GInd}(1,10) = 10$

$$\mathbf{T}^{\langle 00,10\rangle} = \mathbf{T}^{-} \otimes \mathbf{T}^{0} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \mathbf{r}^{\langle 00,10\rangle} = \mathbf{r}^{-} \otimes \mathbf{r}^{0} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}.$$

$m = 10,\ i = 1 : \mathrm{GInd}(1,10) = 10;\ \mathrm{GInd}(0,10) = 00$

$$\mathbf{T}^{\langle 10,00\rangle} = \mathbf{T}^{+} \otimes \mathbf{T}^{0} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathbf{r}^{\langle 10,00\rangle} = \mathbf{r}^{+} \otimes \mathbf{r}^{0} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

$$\mathbf{I}_1\mathbf{P}_{\mathrm{c}}^{\mathrm{ev}} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0.4 \end{pmatrix} = \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}.$$

Подставим полученные векторы в формулу локального апостериорного вывода и получим $\mathbf{P}_{\mathrm{C}}^{1,a} = \begin{pmatrix} 1 \\ \frac{78}{105} \\ 0.4 \\ \frac{12}{35} \end{pmatrix}$.

Теперь найдем вектор-сепаратор. Алфавит вектора-сепаратора и вектор вероятностей его элементов можно получить следуюшим образом:

$A_{\mathrm{sep}} = A_1 \cap A_2 = \{x_1, x_2\} \cap \{x_1, x_3\} = \{x_1\};$

$$\mathbf{P}_{\mathrm{c}}^{\mathrm{sep}} = \begin{pmatrix} 1 \\ p(x_1) \end{pmatrix}; \mathbf{Q} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix};$$

$$\mathbf{P}_c^{\text{sep}} = \mathbf{Q}\mathbf{P}_c^{1,a} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{78}{105} \\ 0.4 \\ \frac{12}{35} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{78}{105} \end{pmatrix}.$$

Воспользуемся формулой пропагации виртуального свидетельства, полученной выше, но сперва вычислим значения всех её составляющих.

$m = 01, i = 0 : \mathrm{GInd}(0, 01) = 00, \mathrm{GInd}(1, 01) = 01;$

$$\mathbf{T}^{\langle 00,01 \rangle} = \mathbf{T}^0 \otimes \mathbf{T}^- = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \mathbf{r}^{\langle 00,01 \rangle} = \mathbf{r}^0 \otimes \mathbf{r}^- = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix};$$

$m = 01, i = 1 : \mathrm{GInd}(1, 01) = 01, \mathrm{GInd}(0, 01) = 00;$

$$\mathbf{T}^{\langle 01,00 \rangle} = \mathbf{T}^0 \otimes \mathbf{T}^+ = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathbf{r}^{\langle 10,00 \rangle} = \mathbf{r}^0 \otimes \mathbf{r}^+ = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix};$$

$$\mathbf{I}_1 \mathbf{P}_c^{\text{sep}} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{78}{105} \end{pmatrix} = \begin{pmatrix} \frac{27}{105} \\ \frac{78}{105} \end{pmatrix}.$$

Подставим полученные векторы в формулу апостериорных вероятностей, посчитаем и получим $\mathbf{P}_c^{2,a} = \begin{pmatrix} 1 \\ \frac{78}{105} \\ \frac{159}{525} \\ \frac{78}{525} \end{pmatrix}$

Заключение

В статье проводится пошаговое рассмотрение одного из случаев алгоритма глобального апостериорного вывода и пример его работы. Рассмотренные матрично-векторные уравнения создают базу для дальнейшего развития алгоритмов глобального апостериорного вывода, исследования чувствительности и устойчивости предложенной модели. Кроме того, матрично-векторная форма записи упрощает программную реализацию за счет возможности использования соответствующих структур и операций.

Тем не менее, в текущей матрично-векторной интерпретации по-прежнему присутствуют элементы функционального описания, что усложняет исследование данной модели. Поэтому в дальнейшем планируется развить имеющийся матрично-векторный подход, а именно решить следующие задачи: доказать предложенный алгоритм формирования матрицы $\mathbf{Q}$ и заменить функции GInd матрично-векторными вычислениями.

Литература

1. Городецкий В. И. Алгебраические байесовские сети --- новая парадигма экспертных систем // Юбилейный сборник трудов институтов Отделения информатики, вычислительной техники и автоматизации РАН. Т. 2. М.: РАН, 1993. С. 120–141
2. Золотин А. А., Тулупьев А. Л. Матрично-векторные алгоритмы глобального апостериорного вывода в алгебраических байесовских сетях // 2017 XX IEEE International Conference on Soft Computing and Measurements (SCM), (St. Petersburg, 25-27 May 2017). St. Petersburg: IEEE, 2017
3. Золотин А.А., Тулупьев А.Л., Сироткин А.В. Матрично-векторные алгоритмы нормировки для локального апостериорного вывода в алгебраических байесовских сетях // Научно-технический вестник информационных технологий, механики и оптики. 2015. Том 15. № 1. С. 78–85
4. Сироткин А.В. Алгебраические байесовские сети: вычислительная сложность алгоритмов логико-вероятностного вывода в условиях неопределенности. Дисс. … канд. физ.-мат. наук. СПб.: СПбГУ, 2011. 218 с.
5. Сироткин А. В., Тулупьев А. Л. Матрично-векторные уравнения локального логико-вероятностного вывода в алгебраических байесовских сетях // Труды СПИИРАН. Вып. 6. --- СПб.: Наука, 2008
6. Тулупьев А. Л. Алгебраические байесовские сети: локальный логико-вероятностный вывод: Учеб. пособие. Элементы мягких

вычислений. СПб.: СПбГУ; ООО Издательство «Анатолия», 2007. 80 с.

7. Тулупьев А.Л., Николенко С.И., Сироткин А.В. Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.
8. Тулупьев А.Л., Сироткин А.В., Золотин А.А. Матричные уравнения нормирующих множителей в локальном апостериорном выводе оценок истинности в алгебраических байесовских сетях // Вестник СПбГУ. Сер. 1. Т. 2 (60). 2015. Вып. 3, С. 379–386
9. Bertone E., Sahin O., Richards R., Roiko A. Extreme events, water quality and health: A participatory Bayesian risk assessment tool for managers of reservoirs //Journal of Cleaner Production. 2016. vol. 135. p. 657-667.
10. Liao Y., Xu B., Wang J.F., Liu X.C. A new method for assessing the risk of infectious disease outbreak //Scientific Reports. 2017. vol. 7.
11. Li S., Zhuang J., Shen S. A three-stage evacuation decision-making and behavior model for the onset of an attack //Transportation Research Part C: Emerging Technologies. 2017. vol. 79. p. 119-135.
12. Marvin H. J. P. et al.Application of Bayesian Networks for hazard ranking of Nanomaterials to Support Human Health Risk Assessment //Nanotoxicology.2017. p. 1-34.
13. Peikert T., Garbe H., Potthast S. Fuzzy-Based Risk Analysis for IT-Systems and Their Infrastructure //IEEE Transactions on Electromagnetic Compatibility. 2017. vol. 59. №. 4. p. 1294-1301.
14. Wang G., Xu T.H., Tang T., Yuan T.M., Wang H.F. A Bayesian network model for prediction of weather-related failures in railway turnout systems //Expert Systems with Applications. 2017. vol. 69. p. 247-256.