

МАТРИЧНЫЕ УРАВНЕНИЯ НОРМИРУЮЩИХ МНОЖИТЕЛЕЙ В ЛОКАЛЬНОМ АПОСТЕРИОРНОМ ВЫВОДЕ ОЦЕНОК ИСТИННОСТИ В АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЯХ*

А. Л. Тулупьев^{1,2}, А. В. Сироткин^{1,2}, А. А. Золотин¹

¹ Санкт-Петербургский государственный университет,

Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7/9

² Санкт-Петербургский институт информатики и автоматизации РАН,

Российская Федерация, 199178, Санкт-Петербург, 14-я линия, 39

Апостериорный вывод является одним из трех видов вывода, на которых основывается обработка фрагментов знаний с вероятностной неопределенностью в интеллектуальных системах поддержки принятия решений, использующих алгебраические байесовские сети. В данной статье даны ключевые термины и формулировки теорем, описывающие локальный апостериорный вывод в алгебраических байесовских сетях на матрично-векторном языке. Основной результат состоит в том, что удалось построить матричные уравнения для нормирующих множителей, фигурирующих в формулах апостериорных вероятностей пропозиций-квантов и идеалов конъюнктов. Сформированные в целом матричные уравнения локального априорного вывода не только упрощают подготовку спецификаций соответствующих алгоритмов вывода и делают реализацию таковых более прозрачной, но также открывают возможность применения классических математических приемов к анализу свойств результатов вывода. Библиогр. 14 назв. Табл. 1.

Ключевые слова: вероятностная логика, байесовские сети, логико-вероятностный вывод, нормирующие множители, знания с неопределенностью, распространение свидетельств, непротиворечивость.

1. Введение. Алгебраические байесовские сети (АБС) предложены В. И. Городецким [2] как новая парадигма моделирования баз фрагментов знаний с неопределенностью и являются одним из классов вероятностных графических моделей. Существенной особенностью АБС является их структура и используемые оценки истинности [6, 7, 9, 11, 12]. Первичной структурой АБС является набор идеалов конъюнктов со скалярными или интервальными оценками истинности их элементов, причем оба варианта оценок формализованы в рамках подхода Н. Нильссона [14]; также рассматриваются вторичная структура АБС и иные виды глобальных структур, построенных над первичной [8, 9, 11, 12, 13]. Идеалы конъюнктов с оценками истинности рассматриваются как математические модели фрагментов знаний [3, 7, 11, 12], а сама АБС — как математическая модель баз фрагментов знаний с неопределенностью.

Возможность применения АБС в интеллектуальных системах, ориентированных на поддержку принятия решений, существенно опирается на решение серии задач по развитию математического и алгоритмического обеспечения логико-вероятностного вывода в этих сетях [5, 11, 12]. Одной из таких проблем является описание локального апостериорного логико-вероятностного вывода на языке матричных (или матрично-векторных) уравнений, что не только существенно упрощает и делает более контролируемой последующую алгоритмическую реализацию такого вывода, но и позволяет применять в анализе свойств его результатов классический математический инструментарий [4, 5, 12].

*Часть публикуемых материалов получена в рамках проекта, выполненного при финансовой поддержке РФФИ (гранты № 12-01-00945-а, 15-01-09001).

В [10] были предложены матричные уравнения, описывающие локальный апостериорный вывод в АБС, которые, однако, содержали элементы, оставшиеся без подходящей интерпретации на матрично-векторном языке. Цель настоящей работы — развить результаты, полученные в [10], предложив матричные уравнения для вычисления указанных элементов, а именно — нормирующих множителей. Чтобы избежать многочисленных повторов, изложение будет вестись в рамках терминов и обозначений, принятых в [10].

2. Матричные уравнения локального апостериорного вывода. Формальная постановка задачи. Зафиксируем конечное множество атомарных пропозициональных формул (атомов) — алфавит $A = \{x_i\}_{i=0}^{n-1}$. Следует заметить, что индексация атомов задает на них порядок; этот порядок используется далее в работе. Над указанными атомами определены два набора «базовых» пропозициональных формул. Первый набор формул — идеал конъюнктов —

$$\{x_{i_1}x_{i_2}\dots x_{i_k} \mid 0 \leq i_1 < i_2 < \dots < i_k \leq n-1, 0 \leq k \leq n\},$$

где $x_{i_1}x_{i_2}\dots x_{i_k}$ означает конъюнкцию соответствующих переменных; сам знак конъюнкции мы для удобства опустим. Каждому из конъюнктов вида $x_{i_1}x_{i_2}\dots x_{i_k}$ можно сопоставить число $2^{i_1} + 2^{i_2} + \dots + 2^{i_k}$ — номер конъюнкта. В определении второго набора формул — множества квантов — используется понятие литерала \tilde{x}_i : на его месте в пропозициональной формуле может стоять либо x_i , либо его отрицание \bar{x}_i . Тогда множество квантов над алфавитом имеет вид $A = \{x_i\}_{i=0}^{n-1} - Q = \{\tilde{x}_0\tilde{x}_1\dots\tilde{x}_{n-1}\}$. Иными словами, квант — это конъюнкция [8, 12], которая для любой переменной из алфавита содержит либо ее саму, либо ее отрицание.

Вслед определяются векторы вероятностей квантов и конъюнктов:

$$\mathbf{P}_c = \begin{pmatrix} 1 \\ p(c_1) \\ \vdots \\ p(c_{2^n-1}) \end{pmatrix} \text{ и } \mathbf{P}_q = \begin{pmatrix} p(q_0) \\ p(q_1) \\ \vdots \\ p(q_{2^n-1}) \end{pmatrix}.$$

Заметим, что вероятности квантов и вероятности конъюнктов выражаются друг через друга с помощью следующих соотношений:

$$\mathbf{P}_q = \mathbf{I}_n \mathbf{P}_c, \quad \mathbf{P}_c = \mathbf{J}_n \mathbf{P}_q, \quad (1)$$

где $\mathbf{I}_n = \mathbf{I}_1^{[n]}$, а $\mathbf{I}_1 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, $\mathbf{J}_n = \mathbf{J}_1^{[n]}$, а $\mathbf{J}_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, причем $[n]$ — степень Кронекера. Ассоциированная с ней операция прямого (тензорного) умножения матриц обозначается как \otimes [1].

Для описания апостериорного вывода мы воспользуемся логико-вероятностной моделью свидетельств в теории АБС. Под свидетельством мы понимаем новые «обусловливающие» данные, которые поступили во фрагмент знаний и с учетом которых нам требуется пересмотреть все (или некоторые) оценки. В данной статье мы рассмотрим только детерминированное свидетельство; другие случаи могут быть сведены к пропагации серии детерминированных свидетельств [5, 6, 12]. Мы говорим, что на вход системы поступило детерминированное свидетельство, если новые сведения представимы в виде конъюнкции атомарных переменных и их отрицаний. Примерами

таких свидетельств могут быть $\langle x_1 \rangle, \langle \bar{x}_0 x_1 \bar{x}_2 \rangle$. Заметим, что такое свидетельство можно разбить на «положительный» и «отрицательный» конъюнкты. В первый входят все положительно означенные атомарные переменные свидетельства, а во второй — отрицательно. При этом и положительной и отрицательной части можно сопоставить индекс, соответствующий десятичному представлению двоичного числа, являющегося характеристическим вектором положительного (или отрицательного) свидетельства, и наши свидетельства можно будет записать следующими эквивалентными обозначениями:

$$\langle \bar{x}_0 x_1 \bar{x}_2 \rangle = \langle x_1, \bar{x}_0 \bar{x}_2 \rangle = \langle 2, 5 \rangle = \langle 010_2, 101_2 \rangle.$$

Далее, мы полагаем, что нам поступило свидетельство $\langle i; j \rangle$. Согласно [10] сопоставим ему матрицу

$$\mathbf{H}^{\langle i; j \rangle} = \tilde{\mathbf{H}}_{n-1}^{\langle i; j \rangle} \otimes \tilde{\mathbf{H}}_{n-2}^{\langle i; j \rangle} \otimes \dots \otimes \tilde{\mathbf{H}}_0^{\langle i; j \rangle},$$

где

$$\tilde{\mathbf{H}}_k^{\langle i; j \rangle} = \begin{cases} \mathbf{H}^+, & \text{если } x_k \text{ входит в } c_i; \\ \mathbf{H}^-, & \text{если } x_k \text{ входит в } c_j; \\ \mathbf{H}^o, & \text{иначе;} \end{cases}$$

причем $\mathbf{H}^+ = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$; $\mathbf{H}^- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$; $\mathbf{H}^o = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

В [10] уравнения локального апостериорного вывода оформлены в виде двух теорем.

Теорема 1. *Вектор вероятностей квантов нормированного апостериорного вывода может быть вычислен по следующей формуле:*

$$\mathbf{P}_q^{\langle i; j \rangle} = \frac{1}{(\mathbf{1}, \mathbf{H}^{\langle i; j \rangle} \mathbf{P}_q)} \cdot \mathbf{H}^{\langle i; j \rangle} \mathbf{P}_q.$$

Для того чтобы перейти к условным вероятностям конъюнктов, нам необходимо умножить полученный вектор на \mathbf{J}_n . Кроме того, подставим в нашу формулу известное нам соотношение (1). Обозначив вектор условных вероятностей конъюнктов через $\mathbf{P}_c^{\langle i; j \rangle}$, получим

$$\mathbf{P}_c^{\langle i; j \rangle} = \mathbf{J}_n \frac{1}{(\mathbf{1}, \mathbf{H}^{\langle i; j \rangle} \mathbf{P}_q)} \cdot \mathbf{H}^{\langle i; j \rangle} \mathbf{I}_n \mathbf{P}_c.$$

Введем обозначение $\mathbf{T}^{\langle i; j \rangle} = \mathbf{J}_n \mathbf{H}^{\langle i; j \rangle} \mathbf{I}_n$, тогда верна следующая теорема.

Теорема 2. *Вектор вероятностей конъюнктов нормированного апостериорного вывода может быть вычислен по следующей формуле:*

$$\mathbf{P}_c^{\langle i; j \rangle} = \frac{1}{(\mathbf{T}^{\langle i; j \rangle} \mathbf{P}_c)[0]} \cdot \mathbf{T}^{\langle i; j \rangle} \mathbf{P}_c;$$

здесь матрица $\mathbf{T}^{\langle i; j \rangle}$ оператора ненормированного апостериорного вывода может быть вычислена по формуле $\mathbf{T}^{\langle i; j \rangle} = \tilde{\mathbf{T}}_{n-1}^{\langle i; j \rangle} \otimes \tilde{\mathbf{T}}_{n-2}^{\langle i; j \rangle} \otimes \dots \otimes \tilde{\mathbf{T}}_0^{\langle i; j \rangle}$, где

$$\tilde{\mathbf{T}}_k^{\langle i; j \rangle} = \begin{cases} \mathbf{T}^+, & \text{если } x_k \text{ входит в } c_i; \\ \mathbf{T}^-, & \text{если } x_k \text{ входит в } c_j; \\ \mathbf{T}^o, & \text{иначе;} \end{cases}$$

причём $\mathbf{T}^+ = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$, $\mathbf{T}^- = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{T}^\circ = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, а $(\mathbf{T}^{(i,j)} \mathbf{P}_c)[0]$ указывает на верхний элемент вектора-результата операции умножения [10].

В [10] нормирующие множители — величины $(\mathbf{1}, \mathbf{H}^{(i,j)} \mathbf{P}_q)$ и $(\mathbf{T}^{(i,j)} \mathbf{P}_c)[0]$ — оставлены без анализа, хотя они являются ключевыми в завершении расчетов результатов апостериорного вывода. Цель статьи будет достигнута, если для расчета указанных величин будут решены задачи по построению соответствующих матричных уравнений.

3. Нормирующие множители. Рассмотрим подробнее нормирующий множитель $(\mathbf{1}, \mathbf{H}^{(i,j)} \mathbf{P}_q)$ и сформулируем теорему в отношении подхода к его вычислению.

Теорема 3. $(\mathbf{1}, \mathbf{H}^{(i,j)} \mathbf{P}_q) = (\mathbf{s}^{(i,j)}, \mathbf{P}_q)$, где $\mathbf{s}^{(i,j)} = \bigotimes_{k=0}^{k=n-1} \tilde{\mathbf{s}}_k^{(i,j)}$, причём $\mathbf{s}^+ = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\mathbf{s}^- = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{s}^\circ = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Доказательство. Рассмотрим скалярное произведение $(\mathbf{1}, \mathbf{H}^{(i,j)} \mathbf{P}_q)$:

$$\begin{aligned} (\mathbf{1}, \mathbf{H}^{(i,j)} \mathbf{P}_q) &= (\mathbf{H}^{(i,j)\text{T}} \mathbf{1}, \mathbf{P}_q) = \left(\left(\bigotimes_{k=0}^{k=n-1} \tilde{\mathbf{H}}_k^{(i,j)\text{T}} \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix}^{[n]}, \mathbf{P}_q \right) = \\ &= \left(\bigotimes_{k=0}^{k=n-1} \left(\tilde{\mathbf{H}}_k^{(i,j)\text{T}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right), \mathbf{P}_q \right). \end{aligned}$$

Матрица $\tilde{\mathbf{H}}_k^{(i,j)}$ имеет лишь три возможных варианта, а значит вектор-селектор $\mathbf{s}^{(i,j)}$ может строиться из двумерных векторов, которые тоже могут иметь только три возможных варианта, так как

$$\tilde{\mathbf{s}}_k^{(i,j)} = \tilde{\mathbf{H}}_k^{(i,j)\text{T}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

причем $\tilde{\mathbf{H}}_k^{(i,j)\text{T}} = \tilde{\mathbf{H}}_k^{(i,j)}$, что относит эти матрицы к классу симметрических. Аналогично $\mathbf{H}^{(i,j)} = \mathbf{H}^{(i,j)\text{T}}$, что также относит матрицу $\mathbf{H}^{(i,j)}$ к классу симметрических матриц. Возможные варианты значений матрицы $\mathbf{H}^{(i,j)}$ приведены выше в перед формулировкой теоремы 1. Введем новые обозначения:

$$\begin{aligned} \mathbf{s}^+ &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{s}^- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \mathbf{s}^\circ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \end{aligned}$$

Используя полученные обозначения и вводя новые, продолжим цепочку равенств:

$$\left(\bigotimes_{k=0}^{k=n-1} \left(\tilde{\mathbf{H}}_k^{(i,j)\text{T}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right), \mathbf{P}_q \right) = \left(\bigotimes_{k=0}^{k=n-1} \tilde{\mathbf{s}}_k^{(i,j)}, \mathbf{P}_q \right) = (\mathbf{s}^{(i,j)}, \mathbf{P}_q) \blacksquare.$$

По построению вектор $\tilde{\mathbf{s}}_k^{(i,j)}$ состоит из 0 и 1 и может интерпретироваться как характеристический вектор или вектор-селектор, соответствующий заданному свидетельству: этот характеристический вектор позволяет отобрать те кванты, которые не противоречат поступившему детерминированному свидетельству $\langle i, j \rangle$. С вычислительной точки зрения интересно то, что с помощью побитовых операций, которые

поддерживаются значительным числом современных языков программирования, в том числе T-SQL, каждый компонент вектора-селектора $\mathbf{s}^{(i,j)}$ может быть выражен непосредственно через индексы $\langle i, j \rangle$, задающие свидетельство.

Пользуясь теоремой 3 решим поставленную задачу и в отношении нормирующего множителя $(\mathbf{T}^{(i,j)} \mathbf{P}_c) [0]$.

Теорема 4. $(\mathbf{T}^{(i,j)} \mathbf{P}_c) [0] = (\mathbf{r}^{(i,j)}, \mathbf{P}_c)$, где $\mathbf{r}^{(i,j)} = \bigotimes_{k=0}^{k=n-1} \tilde{\mathbf{r}}_k^{(i,j)}$, причем

$$\mathbf{r}^+ = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \mathbf{r}^- = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \mathbf{r}^\circ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Доказательство. Примем во внимание, что $(\mathbf{T}^{(i,j)} \mathbf{P}_c) [0] = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}^{[n]}, \mathbf{T}^{(i,j)} \mathbf{P}_c \right)$.

Воспользуемся имеющим место соотношением $\mathbf{T}^{(i,j)} = \mathbf{J}_n \mathbf{H}^{(i,j)} \mathbf{I}_n = \mathbf{J}_1^{[n]} \mathbf{H}^{(i,j)} \mathbf{I}_1^{[n]}$. В результате операции транспонирования получим

$$\mathbf{T}^{(i,j)\top} = \mathbf{I}_1^{[n]\top} \mathbf{H}^{(i,j)\top} \mathbf{J}_1^{[n]\top} = \mathbf{I}_1^{\top[n]} \mathbf{H}^{(i,j)\top} \mathbf{J}_1^{\top[n]}.$$

При таком представлении получается

$$\begin{aligned} \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}^{[n]}, \mathbf{T}^{(i,j)} \mathbf{P}_c \right) &= \left(\mathbf{T}^{(i,j)\top} \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{[n]}, \mathbf{P}_c \right) = \left(\mathbf{I}_1^{\top[n]} \mathbf{H}^{(i,j)\top} \mathbf{J}_1^{\top[n]} \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{[n]}, \mathbf{P}_c \right) = \\ &= \left(\bigotimes_{k=0}^{k=n-1} \left(\mathbf{I}_1^{\top} \tilde{\mathbf{H}}_k^{(i,j)\top} \mathbf{J}_1^{\top} \right) \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{[n]}, \mathbf{P}_c \right) = \\ &= \left(\bigotimes_{k=0}^{k=n-1} \left(\mathbf{I}_1^{\top} \tilde{\mathbf{H}}_k^{(i,j)} \mathbf{J}_1^{\top} \right) \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{[n]}, \mathbf{P}_c \right) = \\ &= \left(\bigotimes_{k=0}^{k=n-1} \left(\mathbf{I}_1^{\top} \tilde{\mathbf{H}}_k^{(i,j)} \mathbf{J}_1^{\top} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right), \mathbf{P}_c \right). \end{aligned}$$

Введем в рассмотрение векторы $\tilde{\mathbf{r}}_k^{(i,j)}$ и построим вектор $\mathbf{r}^{(i,j)} = \bigotimes_{k=0}^{k=n-1} \tilde{\mathbf{r}}_k^{(i,j)}$, где $\tilde{\mathbf{r}}_k^{(i,j)} = \mathbf{I}_1^{\top} \tilde{\mathbf{H}}_k^{(i,j)} \mathbf{J}_1^{\top} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Так как существует всего три варианта матрицы $\tilde{\mathbf{H}}_k^{(i,j)}$, вектор $\tilde{\mathbf{r}}_k^{(i,j)}$ также будет иметь только три варианта. Рассмотрим каждый:

$$\begin{aligned} \mathbf{r}^+ &= \mathbf{I}_1^{\top} \mathbf{H}^+ \mathbf{J}_1^{\top} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \\ \mathbf{r}^- &= \mathbf{I}_1^{\top} \mathbf{H}^- \mathbf{J}_1^{\top} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}; \\ \mathbf{r}^\circ &= \mathbf{I}_1^{\top} \mathbf{H}^\circ \mathbf{J}_1^{\top} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

В системе ранее принятых и введенных обозначений и соглашений получим окончательно

$$\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}^{[n]}, \mathbf{T}^{(i,j)} \mathbf{P}_c \right) = \left(\bigotimes_{k=0}^{k=n-1} \tilde{\mathbf{r}}_k^{(i,j)}, \mathbf{P}_c \right) \blacksquare.$$

Полученные результаты иллюстрируется небольшим числом примеров в таблице.

Векторы $\mathbf{s}^{(i,j)}$ и $\mathbf{r}^{(i,j)}$

Алфавит	Свидетельство	$\mathbf{s}^{(i,j)}$	$\mathbf{r}^{(i,j)}$
$\{x_1, x_0\}$	$x_0 = \langle 1; 0 \rangle = \langle 01; 00 \rangle$	$\mathbf{s}^{(1,0)} = \mathbf{s}^+ \otimes \mathbf{s}^\circ$ $\mathbf{s}^{(1,0)T} = (0 \ 1 \ 0 \ 1)$	$\mathbf{r}^{(1,0)} = \mathbf{r}^+ \otimes \mathbf{r}^\circ$ $\mathbf{r}^{(1,0)T} = (0 \ 1 \ 0 \ 0)$
$\{x_1, x_0\}$	$\bar{x}_0 = \langle 0; 1 \rangle = \langle 00; 01 \rangle$	$\mathbf{s}^{(0,1)} = \mathbf{s}^- \otimes \mathbf{s}^\circ$ $\mathbf{s}^{(0,1)T} = (1 \ 0 \ 1 \ 0)$	$\mathbf{r}^{(0,1)} = \mathbf{r}^- \otimes \mathbf{r}^\circ$ $\mathbf{r}^{(0,1)T} = (1 \ \bar{1} \ 0 \ 0)$
$\{x_2, x_1, x_0\}$	$x_0 = \langle 1; 0 \rangle = \langle 001; 000 \rangle$	$\mathbf{s}^{(1,0)} = \mathbf{s}^+ \otimes \mathbf{s}^{\circ[2]}$ $\mathbf{s}^{(1,0)T} = (0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1)$	$\mathbf{r}^{(1,0)} = \mathbf{r}^+ \otimes \mathbf{r}^{\circ[2]}$ $\mathbf{r}^{(1,0)T} = (0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)$
$\{x_2, x_1, x_0\}$	$x_0 x_1 = \langle 3; 0 \rangle = \langle 011; 000 \rangle$	$\mathbf{s}^{(3,0)} = \mathbf{s}^{+[2]} \otimes \mathbf{s}^\circ$ $\mathbf{s}^{(3,0)T} = (0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 1)$	$\mathbf{r}^{(3,0)} = \mathbf{r}^{+[2]} \otimes \mathbf{r}^\circ$ $\mathbf{r}^{(3,0)T} = (0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0)$
$\{x_2, x_1, x_0\}$	$x_0 \bar{x}_1 = \langle 1; 2 \rangle = \langle 001; 010 \rangle$	$\mathbf{s}^{(1,2)} = \mathbf{s}^+ \otimes \mathbf{s}^- \otimes \mathbf{s}^\circ$ $\mathbf{s}^{(1,2)T} = (0 \ 1 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0)$	$\mathbf{r}^{(1,2)} = \mathbf{r}^+ \otimes \mathbf{r}^- \otimes \mathbf{r}^\circ$ $\mathbf{r}^{(1,2)T} = (0 \ 1 \ 0 \ \bar{1} \ 0 \ 0 \ 0 \ 0)$

Примечание. В таблице -1 для удобства обозначается как $\bar{1}$.

Вектор $\mathbf{r}^{(i,j)}$ удобно рассматривать как вектор-редистрибутор, отвечающий за перераспределение вклада каждого конъюкта в нормирующий множитель. Так же как и в случае вектора-селектора, компоненты вектора $\mathbf{r}^{(i,j)}$ можно вычислять с помощью побитовых логических операций через индексы i, j , задающие свидетельство. Это позволяет не формировать сразу весь вектор размерности 2^n в памяти ЭВМ, а отложить вычисления до тех пор, пока в них не потребуется конкретная его компонента.

Кратко отметим, что выведенные матричные уравнения открывают возможность использовать классические математические приемы в анализе чувствительности обобщающих нормирующих множителей к вариациям исходных данных. Анализ чувствительности сведется, в зависимости от выбранных метрик и ограничений, к исследованию следующих конструкций:

$$\left(\mathbf{r}^{(i,j)}, \mathbf{P}_c \right) - \left(\mathbf{r}^{(i,j)}, \hat{\mathbf{P}}_c \right) = \left(\mathbf{r}^{(i,j)}, \mathbf{P}_c - \hat{\mathbf{P}}_c \right) = \left(\mathbf{r}^{(i,j)}, \Delta \mathbf{P}_c \right), \quad (2)$$

$$\left(\mathbf{s}^{(i,j)}, \mathbf{P}_q \right) - \left(\mathbf{s}^{(i,j)}, \hat{\mathbf{P}}_q \right) = \left(\mathbf{s}^{(i,j)}, \mathbf{P}_q - \hat{\mathbf{P}}_q \right) = \left(\mathbf{s}^{(i,j)}, \Delta \mathbf{P}_q \right), \quad (3)$$

где $\Delta \mathbf{P}_q$ и $\Delta \mathbf{P}_c$ описывают рассматриваемую допустимую вариацию исходных данных.

Поясним применение термина *допустимая* вариация данных: он означает, что не только исходные данные \mathbf{P}_c и \mathbf{P}_q отвечают требованиям непротиворечивости [6, 12], но и векторы с вариацией значений элементов $\hat{\mathbf{P}}_c$ и $\hat{\mathbf{P}}_q$ должны отвечать тем же требованиям, что может быть оформлено матричными соотношениями вида [6, 10]

$$\mathbf{I}_n \mathbf{P}_c \geq \mathbf{0}; \quad \mathbf{P}_q \geq \mathbf{0}, \quad (\mathbf{P}_q, \mathbf{1}) = 1; \quad \mathbf{I}_n \hat{\mathbf{P}}_c \geq \mathbf{0}; \quad \hat{\mathbf{P}}_q \geq \mathbf{0}, \quad (\hat{\mathbf{P}}_q, \mathbf{1}) = 1.$$

Знание структуры левых множителей и состава их элементов (0, 1 и -1 , либо 0 и 1) в скалярных произведениях из соотношений (2)–(3) открывает богатые возможности для построения оценок «сверху» вариации результатов, а также может оказаться полезным для построения оценок «снизу». Наконец, полученные векторы $\mathbf{r}^{(i,j)}$ и $\mathbf{s}^{(i,j)}$ также могут быть использованы в анализе чувствительности результатов локального апостериорного вывода.

4. Заключение. В статье предложено развитие подхода к описанию процессов локального логико-вероятностного вывода в алгебраических байесовских сетях: за счет углубления анализа структуры нормирующих множителей удалось вывести матричные уравнения для их расчета. Такие уравнения не только упрощают подготовку спецификаций соответствующих алгоритмов вывода и делают реализацию таковых более прозрачной, но также открывают возможность применения классических математических приемов к анализу свойств результатов вывода.

Литература

1. Беллман Р. Введение в теорию матриц. М.: Наука, Гл. ред. физ.-мат. лит-ры, 1969. 368 с.
2. 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.
3. Городецкий В. И., Тулупьев А. Л. Формирование непротиворечивых баз знаний с неопределенностью // Изв. РАН. Сер. Теория и системы управления. 1997. Т. 5. С. 33—42.
4. Сироткин А. В. Вычислительная сложность алгоритмов локального апостериорного вывода в алгебраических байесовских сетях // Труды СПИИРАН. 2011. № 3(18). С. 188—214.
5. Сироткин А. В., Тулупьев А. Л. Локальный априорный вывод в алгебраических байесовских сетях: комплекс основных алгоритмов // Труды СПИИРАН. Вып. 5. 2007. С. 100—111.
6. Тулупьев А. Л. Алгебраические байесовские сети: локальный логико-вероятностный вывод: Учеб. пособие. Элементы мягких вычислений. СПб.: ООО Издательство «Анатолия», 2007. 80 с.
7. Тулупьев А. Л. Алгебраические байесовские сети: логико-вероятностный подход к моделированию баз знаний с неопределенностью. СПб.: СПИИРАН, 2000. 282 с.
8. Тулупьев А. Л. Байесовские сети: логико-вероятностный вывод в циклах. СПб.: Изд-во С.-Петербург. ун-та, 2008. 140 с.
9. Тулупьев А. Л., Сироткин А. В. Алгебраические байесовские сети: принцип декомпозиции и логико-вероятностный вывод в условиях неопределенности // Информационно-измерительные и управляющие системы. 2008. Т. 6, № 10. С. 85—87.
10. Тулупьев А. Л., Сироткин А. В. Матричные уравнения локального логико-вероятностного вывода оценок истинности элементов в алгебраических байесовских сетях // Вестн. С.-Петербург. ун-та. Сер. 1. 2012. Вып. 3. С. 63—72.
11. Тулупьев А. Л., Николенко С. И., Сироткин А. В. Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.
12. Тулупьев А. Л., Сироткин А. В., Николенко С. И. Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах. СПб.: Изд-во С.-Петербург. ун-та, 2009. 400 с.
13. Фильченков А. А. Иерархия глобальных структур алгебраической байесовской сети как система графов и гиперграфов // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 1. С. 75—81.
14. Nilsson N. J. Probabilistic Logic // Artificial Intelligence. 1986. Vol. 47. Amsterdam: Elsevier Science Publishers B.V., 1986. P. 71—87.

Статья поступила в редакцию 26 марта 2015 г.

Сведения об авторах

Тулупьев Александр Львович — доктор физико-математических наук, профессор;
alexander.tulupyev@gmail.com

Сироткин Александр Владимирович — кандидат физико-математических наук, доцент;
alexander.sirotkin@gmail.com

Золотин Андрей Алексеевич — студент; andrey.zolotin@gmail.com

MATRIX EQUATIONS FOR NORMALIZING FACTORS IN THE LOCAL POSTERIORI INFERENCE OF THE TRUTH ESTIMATION IN ALGEBRAIC BAYESIAN NETWORKS

Alexander L. Tulupyev^{1,2}, Alexander V. Sirotkin^{1,2}, Andrey A. Zolotin¹

¹ St.Petersburg State University, Universitetskaya nab., 7/9, St.Petersburg, 199034, Russian Federation; alexander.tulupyev@gmail.com, alexander.sirotkin@gmail.com, andrey.zolotin@gmail.com

² St.Petersburg Institute for Informatics and Automation RAS,

14-ya liniya V. O., 39, St.Petersburg, 199178, Russian Federation; alexander.tulupyev@gmail.com

A posteriori inference is one of three kinds of probabilistic-logic inferences, which the processing of fragments of knowledge with probabilistic uncertainty using Bayesian networks is based on. In this paper, the key terms and theorems statements describing local posteriori inference in algebraic Bayesian networks are given in terms of matrix-vector language. The main result is that we managed to construct a matrix-vector equation for the normalizing factors, appearing in the formulas of posterior probabilities of quantum and conjuncts ideals. The whole local posteriori inference equations presented in matrix-vector terms simplify specifications of related inference algorithms and make their implementation more transparent as well as open a way to the classical mathematical technique usage for the inference results sensitivity analysis. Refs 14. Tables 1.

Keywords: probabilistic logic, Bayesian networks, probabilistic-logic inference, normalizing factors, uncertain knowledge, evidence propagation, consistency.

References

1. Bellman R., *Introduction to Matrix Analysis* (McGraw-Hill, New York, 1960; Nauka, Moscow, 1969).
2. Gorodetskii V.I., "Algebraic Bayesian networks—new paradigm of expert systems", *Anniversary Collection of Proceedings of Institutes of the Russian Academy of Sciences Department of Informatics, Computer Science and Automation* **2** (Ross. Akad. Nauk, Moscow, 1993) [in Russian].
3. Gorodetskii V.I., Tulup'ev A.L., "Formation of noncontradicting knowledge bases with an uncertainty", *Izv. Ross. Akad. Nauk, Ser. Teor. Sist. Upr.* **5**, 33–42 (1997). MathSciNet
4. Sirotkin A.V., "Computational complexity of local posteriori inference algorithms in algebraic Bayesian networks", *Tr. SPIIRAN* no.3(18), 188–214 (2011).
5. Sirotkin A.V., Tulup'ev A.L., "Local a priori inference in algebraic Bayesian network: A set of basic algorithms", *Tr. SPIIRAN* no.5, 100–111 (2007).
6. Tulup'ev A.L., "Algebraic Bayesian networks: a local probabilistic logic inference", *Handbook on Elements of Soft Computing* (Anatoliya, St. Petersburg, 2007) [in Russian].
7. Tulup'ev A.L., *Algebraic Bayesian Networks: a Probabilistic Logic Approach to the Modelling of Knowledge Bases with an Uncertainty* (SPIIRAN, St. Petersburg, 2000) [in Russian].
8. Tulup'ev A.L., *Bayesian Networks: A Probabilistic Logic Inference in Cycles* (St. Peterb. Gos. Univ., St. Petersburg, 2008) [in Russian].
9. Tulup'ev A.L., Sirotkin A.V., "Algebraic Bayesian networks: Principle of decomposition and probabilistic logic inference under uncertainty", *Inf.-Izmer. Upravl. Sist.* **6**(10), 85–87 (2008).
10. Tulup'ev A.L., Sirotkin A.V., "Matrix equations for local probabilistic logic inference of truth values of elements in algebraic Bayesian networks", *Vestn. St. Peterb. Univ., Ser. 1: Mat., Mekh., Astron.* no.3, 63–72 (2012).
11. Tulup'ev A.L., Nikolenko S.I., Sirotkin A.V., *Bayesian Networks: A Probabilistic Logic Approach* (Nauka, St. Petersburg, 2006) [in Russian].
12. Tulup'ev A.L., Sirotkin A.V., Nikolenko S.I., *Bayesian Belief Networks: Probabilistic Logic Inference in Acyclic Directed Graphs* (St. Peterb. Gos. Univ., St. Petersburg, 2009) [in Russian].
13. Fil'chenkov A.A., "Hierarchy of algebraic Bayesian network global structures as a system of graphs and hypergraphs", *Nauchno-Tekhn. Vestn. Inf. Tekhnol., Mekh. Opt.* **13**(1), 75–81 (2013).
14. Nilsson N.J., "Probabilistic Logic", *Artificial Intelligence* **47** (Elsevier Science Publishers B.V., Amsterdam, 1986, 71–87).