МАТРИЧНО-ВЕКТОРНЫЕ УРАВНЕНИЯ ЛОКАЛЬНОГО ЛОГИКО-ВЕРОЯТНОСТНОГО ВЫВОДА В АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЯХ*

А. В. Сироткин, А. Л. Тулупьев

Санкт-Петербургский институт информатики и автоматизации РАН

СПИИРАН, 14-я линия ВО, д. 39, Санкт-Петербург, 199178

<avs@iias.spb.su>, <alt@iias.spb.su>

УДК 004.8

Сироткин А. В., Тулупьев А. Л. **Матрично-векторные уравнения локального логико-вероятностного вывода в алгебраических байесовских сетях** // Труды СПИИРАН. Вып. 6. — СПб.: Наука, 2008.

Аннотация. *Обработка фрагментов знаний с вероятностной неопределенностью в интеллектуальных системах поддержки принятия решений включает три вида локального логико-вероятностного вывода: проверку и поддержание непротиворечивости, априорный и апостериорный выводы. В настоящей статье вычислительные формулы, на которые опираются перечисленные виды логико-вероятностного вывода, представлены на матрично-векторном языке, что сокращает, упрощает и делает более строгой последующую спецификацию алгоритмов вывода. Кроме того, использование матрично-векторного языка открывает новые возможности для исследования свойств результатов локального логико-вероятностного вывода.* — Библ. 8 назв.

UDC 004.8

Sirotkin A.V., Tulupyev A. L. **Matrix-Vector Equations for Local Probabilistic Logic Inference in Algebraic Bayesian Network** // SPIIRAS Proceedings. Issue 6. — SPb.: Nauka, 2008.

Abstract. *Probabilistically uncertain knowledge patterns processing in intellectual decision support systems falls into three kinds of probabilistic-logic inference, such as reconciliation, a priori and a posteriori inference. The paper presents formulae that allow for putting the process down in terms of matrix-vector language.* — Bibl. 8 Items.

1. Введение

Существуют различные подходы к построению баз фрагментов знаний с вероятностной неопределенностью для интеллектуальных систем для поддержки принятия решений. Одним из подходов является парадигма алгебраических байесовских сетей, предложенная В. И. Городецким в 1983 году [1, 2]. В теории алгебраических байесовских сетей (АБС) математической моделью фрагмента знаний (ФЗ) является идеал цепочек конъюнкций (или идеал конъюнктов, который будет определен далее) с оценками вероятности истинности входящих в него элементов. Для формализации понятия вероятности истинности пропозициональной формулы используется способ, основанный на подходе Н. Нильссона [3].

Цель настоящей работы — дать на матрично-векторном языке формальное описание локального логико-вероятностного вывода в ФЗ АБС.

* Исследования, результаты которых представлены в настоящей работе, были частично поддержаны грантом по Программе «Участник молодежного научно-инновационного конкурса» (УМНИК) договор № У-2007-3/5-1 в 2007-2008 гг. и грантами Фонда содействия отечественным учёным по программе «Лучшие аспиранты РАН» за 2008 год и «Выдающиеся ученые. Кандидаты и доктора наук РАН» за 2004–2007 гг.

2. Базовые объекты и перенумерация

В первую очередь зафиксируем конечное множество атомарных пропозициональных формул (атомов) — алфавит $A=\{x_i\}_{i=0}^{n-1}$. Отметим, что для удобства мы будем вести нумерацию переменных с нуля. Определим над указанными атомами два набора «базовых» пропозициональных формул.

Первый набор формул — идеал цепочек конъюнкций (идеал конъюнктов)

$$\left\{x_{i_1}x_{i_2}\ldots x_{i_k} \mid 0\le i_1<i_2<\ldots<i_k\le n-1, k\le n\right\},$$

где $x_{i_1}x_{i_2}\ldots x_{i_k}$ означает конъюнкцию соответствующих переменных; сам знак конъюнкции мы для удобства будем опускать. Каждому из конъюнктов вида $x_{i_1}x_{i_2}\ldots x_{i_k}$ можно сопоставить число $2^{i_1}+2^{i_2}+\ldots+2^{i_k}$ — номер конъюнкта. Более того, если представить полученное число в двоичной записи, то переменная x_i будет входить в конъюнкт тогда и только тогда, когда i-й бит номера будет равен единице (нумеровать биты предлагается начиная с младшего разряда, считая его нулевым битом).

Для определения второго набора формул — множества квантов будет полезным следующее обозначение. Литерал (аргументное место) $\tilde{x}_i$ обозначает, что на его месте в формуле может стоять либо x_i, либо его отрицание $\overline{x}_i$. Тогда множество квантов над алфавитом $A=\{x_i\}_{i=0}^{n-1}$ — $Q=\{\tilde{x}_0\tilde{x}_1\ldots\tilde{x}_{n-1}\}$. Иными словами, квант — это конъюнкция, которая для любой переменной содержит либо ее саму, либо ее отрицание. Для нумерации квантов мы воспользуемся способом, аналогичным нумерации квантов. Выделим «положительную» часть конъюнкта (множество положительно означенных переменных) и рассмотрим ее как конъюнкт. Номер этого конъюнкта и будет номером рассматриваемого кванта. Таким образом, единице в двоичной записи номера соответствует положительное вхождение переменной, а нулю — отрицательное, при этом рассматриваются все n битов (то есть с учетом лидирующих нулей).

После введения нумерации квантов и конъюнктов можно определить векторы вероятностей квантов и конъюнктов:

$$\mathbf{P_c}=\begin{pmatrix}1\\ p(c_1)\\ \vdots\\ p(c_{2^n-1})\end{pmatrix} \text{ и } \mathbf{P_q}=\begin{pmatrix}p(q_0)\\ p(q_1)\\ \vdots\\ p(q_{2^n-1})\end{pmatrix},$$

где c_i — конъюнкт номер i, а q_i — i-й квант. Появление единицы в первом случае вполне оправдано, так как согласно определению c_0 — пустой конъюнкт, соответствующий тождественной истине. После введения перенумерации и определения базовых объектов (конъюнктов и квантов) перейдем к определению вероятности над пропозициональными формулами.

3. Оценки вероятностей над пропозициональными формулами

Пусть $F_0(A)$ — множество всех пропозициональных формул над заданным алфавитом $A=\{x_i\}_{i=1}^{n}$. Определим множество отличающихся формул как фактор-множество всех формул по условию эквивалентности: $F(A)=F_0(A)/\equiv$. Над n атомарными переменными подобных формул существует ровно 2^{2^n}.

По теореме о совершенной нормальной дизъюнктивной форме любая пропозициональная формула может быть представлена в виде дизъюнкции конечного числа *квантов.* Так как при любом зафиксированном означивании всех переменных x_1, x_2, …, x_n никакие два разных кванта не могут быть одновременно истинны, а, с другой стороны, один из них заведомо истинен, можно рассмотреть множество Q как множество элементарных событий. Задав вероятность на квантах, можно в свою очередь построить вероятностное пространство, на котором будет определена вероятность любой пропозициональной формулы. За более подробным описанием аксиоматики вероятностной логики можно обратиться, например, к [3, 4].

Определение вероятности на элементах множества Q потребует следующих ограничений:

$$\forall q\in Q \;\; p(q)\geq 0; \tag{1а}$$

$$\sum_{q\in Q} p(q)=1. \tag{1б}$$

Пользуясь введенными выше векторами, можно переписать эти условия следующим образом:

$$\mathbf{P_q}\geq \mathbf{0}; \tag{2а}$$

$$(\mathbf{1},\mathbf{P_q})=1. \tag{2б}$$

В алгебраических байесовских сетях мы в основном работаем не с квантами, а с конъюнктами. В монографии [4] показано, что через вероятности конъюнктов, как и через вероятности квантов можно выразить вероятность любой пропозициональной формулы. Переход между этими двумя «базисами» можно выразить следующей формулой:

$$\mathbf{P_q}=\mathbf{I}_n\times\mathbf{P_c}. \tag{3}$$

При этом ограничения (1) принимают вид:

$$\mathbf{I}_n\times\mathbf{P_c}\geq\mathbf{0}. \tag{4}$$

Здесь n — число атомов, а матрица $\mathbf{I}_n$ — матрица перехода от вероятностей квантов к вероятностям конъюнктов.

Матрица $\mathbf{I}_n$ имеет очень четкую структуру, которую удобнее всего описать рекуррентно.

$$\mathbf{I}_1=\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix},\; \ldots,\; \mathbf{I}_n=\mathbf{I}_1\otimes\mathbf{I}_{n-1}=\mathbf{I}_1\otimes\mathbf{I}_1^{[n-1]}=\mathbf{I}_1^{[n]};$$

например,

$$\mathbf{I}_2 = \mathbf{I}_1 \otimes \mathbf{I}_1 = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Здесь $\otimes$ означает кронекерово (тензорное) произведение матриц; $\mathbf{I}_n$ — это кронекерова степень матрицы $\mathbf{I}_1$, то есть $\mathbf{I}_n = \mathbf{I}_1^{[n]}$ [5]. Кроме рассмотренной матрицы $\mathbf{I}_n$ будет использоваться и обратная ей — $\mathbf{J}_n$, удовлетворяющая условию:

$$\mathbf{P_c} = \mathbf{J}_n \times \mathbf{P_q}. \tag{5}$$

$$\mathbf{J}_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},\ \mathbf{J}_n = \mathbf{J}_1 \otimes \mathbf{J}_{n-1} = \mathbf{J}_1 \otimes \mathbf{J}_1^{[n-1]} = \mathbf{J}_1^{[n]};$$

$$\mathbf{J}_2 = \mathbf{J}_1 \otimes \mathbf{J}_1 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

4. Непротиворечивость фрагмента знаний

Основным элементом АБС является фрагмент знаний (ФЗ); исследуем вопросы его непротиворечивости. ФЗ представляет с собой идеал конъюнктов с оценками истинности. При этом можно выделить два основных случая:

1. Оценка каждого конъюнкта задана скалярно (точечно); совокупность всех оценок представлена в виде вектора оценок $\mathbf{P_c}$ — тогда для проверки непротиворечивости достаточно проверить выполнение условий (4).

2. Заданы два вектора — верхние и нижние оценки $\mathbf{P^-}$ и $\mathbf{P^+}$; мы говорим, что подобные оценки непротиворечивы, если:

$$\forall i : 1 \le i \le 2^n - 1\ \forall \varepsilon : \mathbf{P^-}[i] \le \varepsilon \le \mathbf{P^+}[i] \quad \exists \mathbf{P_c} : \left(\mathbf{P^-} \le \mathbf{P_c} \le \mathbf{P^+}\right) \& \left(\mathbf{P_c}[i] = \varepsilon\right) \& \left(\mathbf{I}_n \times \mathbf{P_c} \ge \mathbf{0}\right).$$

Здесь запись вида $\mathbf{P_c}[i]$ означает i-й элемент вектора $\mathbf{P_c}$. Если выразить словами, то для любой оценки любого конъюнкта (кроме нулевого, оценка для которого всегда единица), лежащей в границах определенных векторами $\mathbf{P^-}$ и $\mathbf{P^+}$, существует совокупность оценок вероятностей всех остальных конъюнктов, лежащих в границах, определенных $\mathbf{P^-}$ и $\mathbf{P^+}$ и задающих непротиворечивый фрагмент знаний со скалярными оценками.

Как уже было сказано, в первом случае нам достаточно проверить, непротиворечива оценка или нет. В случае с интервальными оценками вопрос не только в том, что непротиворечивы они или нет, но и если противоречивы, то можно ли их сузить до непротиворечивых. Решения этих двух вопросов взаимосвязанны. Для проверки непротиворечивости необходимо решить серию задач линейного программирования (ЗЛП) [4]. Переменными данных задач будут точечные вероятности $\mathbf{P_c}$, а ограничения будут двух типов: $\mathbf{P^-} \le \mathbf{P_c} \le \mathbf{P^+}$ и $\mathbf{I}_n \times \mathbf{P_c} \ge \mathbf{0}$. Осталось определить только целевые функции. Целевыми функциями будут максимизация и минимизация $\mathbf{P_c}[i]$ для каждого $i:1\le i\le 2^n-1$. Решение этой серии ЗЛП позволяет определить, непротиворечив ли ФЗ; в этом случае все ЗЛП будут разрешимы и соответствующие максимумы и минимумы совпадают с заданными границами. Если хотя бы одна из ЗЛП дала результат, отличный от заданных границ, то соответствующие максимумы и минимумы дадут наибольший по включению набор интервальных оценок, задающий непротиворечивый ФЗ и лежащий в указанных границах. А если хоть одна из ЗЛП оказалась неразрешима, то значит такого сужения не существует и ФЗ противоречив.

5. Локальный априорный вывод

Сформировав непротиворечивый ФЗ, перейдем к вопросу об оценивании вероятности произвольной пропозициональной формулы, заданной над теми же атомами, что и фрагмент знаний. Суть априорного вывода — построение оценок для некоторой формулы на основе оценок другого набора формул. В контексте АБС наиболее естественно в качестве набора формул, уже имеющих оценки, рассматривать элементы ФЗ. Мы не будем подробно останавливаться на вопросах обработки формул и перехода к их СДНФ (совершенной дизъюнктивной нормальной форме), за этим можно обратиться к статье [7].

Рассмотрим пропозициональную формулу f, вероятность истинности которой требуется оценить. Обозначим через $\mathbf{L}_f$ вектор, содержащий 2^n элементов, каждый из которых является единицей или нулем в зависимости от того, входит или не входит соответствующий по номеру квант в СДНФ формулы f (нумерация квантов была описана в разд. 2).

Пусть ФЗ — с оценками истинности, то есть мы имеем ограничения вида: $\mathbf{P}^- \le \mathbf{P_c} \le \mathbf{P}^+$, где $\mathbf{P_c}$ — вектор вероятностей конъюнктов входящих в ФЗ, а $\mathbf{P}^-$ и $\mathbf{P}^+$ — векторы, состоящие из соответствующих нижних и верхних оценок.

Вероятность формулы f можно выразить следующим образом:

$$p(f) = (\mathbf{L}_f, \mathbf{P_q}).$$

Для перехода к вероятностям конъюнктов воспользуемся уже применявшейся выше формулой (3) и получим:

$$p(f) = (\mathbf{L}_f, \mathbf{I_n} \times \mathbf{P_c}) = (\mathbf{I_n}^T \times \mathbf{L}_f, \mathbf{P_c}).$$

Таким образом, мы выразили вероятность произвольной формулы через вероятности конъюнктов. Чтобы оценить $p(f)$, построим задачу линейного программирования.

Переменными нашей ЗЛП, как и в случае поддержания непротиворечивости, будут элементы вектора $\mathbf{P_c}$. Ограничениями будут $\mathbf{P}^- \le \mathbf{P_c} \le \mathbf{P}^+$, заданные в ФЗ, и неравенства $\mathbf{I}_n \times \mathbf{P_c} \ge \mathbf{0}$, заданные аксиоматикой теории вероятностей. Целевая функция описывается $p(f) = (\mathbf{I_n}^T \times \mathbf{L}_f, \mathbf{P_c})$. Найдем максимум и минимум целевой функции — это и будут искомые оценки вероятности формулы f. Если же построенная ЗЛП не будет иметь решения, то это означает, что исходный набор оценок был противоречив.

Кроме такого подхода к априорному выводу, можно рассмотреть еще построение (или уточнение) оценок ФЗ на основе оценок произвольного набора

Труды СПИИРАН. 2008. Вып. 6. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2008. Issue 6. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

формул, оно происходит аналогично процессу поддержания непротиворечивости, только в ЗЛП добавляется условие $p_f^- \le (\mathbf{I_n}^T \times \mathbf{L}_f, \mathbf{P_c}) \le p_f^+$, где p_f^- и p_f^+ — нижняя и верхняя оценка вероятности формулы f. Более подробное описание ЗЛП в данном случае можно найти в [4].

6. Перестановки, переозначивания и свидетельства

Для описания последнего вида вывода нам потребуется описать логико-вероятностную модель свидетельств в теории АБС, а также ввести ряд вспомогательных обозначений, которые позволят нам изложить оставшийся материал тоже на матрично-векторном языке.

Под свидетельством мы понимаем «новые» данные, которые поступили во фрагмент знаний и с учетом которых нам требуется пересмотреть все (или некоторые) оценки. Сама суть апостериорного вывода заключается в оценивании условной вероятности элементов ФЗ относительно поступившего свидетельства.

Свидетельства, применяемые в АБС, разделяются на несколько видов:

- детерминированные свидетельства;
- стохастические свидетельства;
- неточные свидетельства.

В данной статье мы рассмотрим только детерминированное свидетельство; оставшиеся два случая могут быть сведены к пропагации серии детерминированных свидетельств [4]. 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 $\langle x_1 \rangle$, $\langle x_2 \bar{x}_3 \rangle$, $\langle x_1 \bar{x}_3 x_4 \rangle$; угловые скобки используются для обозначения того, что соответствующая конъюнкция рассматривается нами как свидетельство. 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 частям можно сопоставить индексы в соответствии с перенумерацией, приведенной во втором разделе настоящей работы, и наши свидетельства можно будет записать следующими эквивалентными обозначениями:

$$\langle x_1 \rangle = \langle x_1 \rangle = \langle 2;0 \rangle = \langle 00010_2;00000_2 \rangle,$$

$$\langle x_2 \bar{x}_3 \rangle = \langle x_2, \bar{x}_3 \rangle = \langle 4;8 \rangle = \langle 00100_2;01000_2 \rangle,$$

$$\langle x_1 \bar{x}_3 x_4 \rangle = \langle x_1 x_4, \bar{x}_3 \rangle = \langle 18;8 \rangle = \langle 10010_2;01000_2 \rangle.$$

Для начала рассмотрим ситуацию, когда детерминированное свидетельство имеет вид $\langle m;0 \rangle$, то есть состоит из конъюнкта. Рассмотрим матрицу $\mathbf{Q}^{\langle m;0 \rangle}$, которая действует на вектор $\mathbf{P_q}$ следующим образом: вероятности квантов, которые согласованы со свидетельством $\langle m;0 \rangle$ (содержат его в себе как подконъюнкт) остаются неизменными, а вероятности всех остальных квантов приравниваются к нулю. Такая матрица состоит из нулей и некоторого количества единиц на главной диагонали:

$$\mathbf{Q}^{\langle m;0\rangle}[i,j]=\begin{cases}\delta(i,j), i\wedge m=m,\\ \quad 0, i\wedge m\neq m.\end{cases}$$

Здесь δ — дельта Кронекера, а конъюнкцию между индексами следует рассматривать как побитовую.

Однако после такого преобразования вектор $\mathbf{Q}^{\langle m;0\rangle}\times\mathbf{P_q}$ не будет удовлетворять требованию (1*б*), поэтому следует рассмотреть нормированный вектор

$$\mathbf{P_q}^{\langle m;0\rangle}=\frac{1}{(\mathbf{1},\mathbf{Q}^{\langle m;0\rangle}\times\mathbf{P_q})}\mathbf{Q}^{\langle m;0\rangle}\times\mathbf{P_q}.$$

В данном случае возможно, что $(\mathbf{1},\mathbf{Q}^{\langle m;0\rangle}\times\mathbf{P_q})=0$, это значит, что вероятность поступившего свидетельства была равна нулю, что в свою очередь значит, что данное свидетельство поступить не могло. В подобной ситуации [6, 7] всем конъюнктам в качестве оценки приписывается интервал $[0;1]$. Отметим, что вектор $\mathbf{P_q}^{\langle m;0\rangle}$ содержит в точности вероятности вида $p\left(\tilde{x}_0\tilde{x}_1\ldots\tilde{x}_{n-1}\,|\,\langle m;0\rangle\right)$. Теперь можно построить вектор

$$\mathbf{P_c}^{\langle m;0\rangle}=\mathbf{J}_n\times\mathbf{P_q}^{\langle m;0\rangle}=\mathbf{J}_n\times\frac{1}{(\mathbf{1},\mathbf{Q}^{\langle m;0\rangle}\times\mathbf{I}_n\times\mathbf{P_c})}\mathbf{Q}^{\langle m;0\rangle}\times\mathbf{I}_n\times\mathbf{P_c},$$

что также даст нам условные вероятности, но уже для конъюнктов, то есть для элементов исходного ФЗ. То, что мы проделали, является частным случаем апостериорного вывода, лежащим в основе реализации других его случаев [6, 7].

Если детерминированное свидетельство содержит «отрицательную» часть, используется особая замена переменных — переозначивание.

Пусть задан непротиворечивый вектор вероятностей $\mathbf{P_c}$ над множеством атомарных пропозиций $\{x_0,x_1,\ldots,x_{n-1}\}$. Как, имея эти данные, выразить через него вектор $\mathbf{P_c}^{\langle -1\rangle}$, который представляет собой вектор вероятностей, аналогичный $\mathbf{P_c}$, но построенный над множеством атомов $\{\bar{x}_0,x_1,\ldots,x_{n-1}\}$? Верхний индекс вида $-k$ мы будем использовать для обозначения атомарных пропозиций, получающих отрицательное означивание, при этом само k — максимальный индекс конъюнкта, состоящий только из отрицательно означенных атомов, входящих в свидетельство. Рассмотрим $\mathbf{P_q}=\mathbf{I}_n\times\mathbf{P_c}$ и $\mathbf{P_q}^{\langle -1\rangle}=\mathbf{I}_n\times\mathbf{P_c}^{\langle -1\rangle}$. Заметим, что векторы $\mathbf{P_q}$ и $\mathbf{P_q}^{\langle -1\rangle}$ отличаются только порядком элементов. Построим соответствующую матрицу переходов

$$\mathbf{A}^{\langle -1\rangle}=\begin{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\ \mathbf{0}&\begin{pmatrix}0&1\\1&0\end{pmatrix}&\mathbf{0}&\mathbf{0}\\ \mathbf{0}&\mathbf{0}&\ddots&\mathbf{0}\\ \mathbf{0}&\mathbf{0}&\mathbf{0}&\begin{pmatrix}0&1\\1&0\end{pmatrix}\end{pmatrix}.$$

Труды СПИИРАН. 2008. Вып. 6. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2008. Issue 6. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

В общем случае матрица $\mathbf{A}^{\langle -k\rangle}$ будет содержать по одной единице в каждой строке, при этом i–я строка будет содержать единицу в столбце с номером $(i\wedge\neg k)\vee(\neg i\wedge k)$, логические операции здесь следует читать как побитовые. Теперь выразим, что требовалось:

$$\mathbf{P}_{\mathbf{c}}^{\langle -1\rangle}=\mathbf{J}_n\times\mathbf{P}_{\mathbf{q}}^{\langle -1\rangle}=\mathbf{J}_n\times(\mathbf{A}^{\langle -1\rangle}\times\mathbf{P}_{\mathbf{q}})=\mathbf{J}_n\times(\mathbf{A}^{\langle -1\rangle}\times(\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}}))=(\mathbf{J}_n\times\mathbf{A}^{\langle -1\rangle}\times\mathbf{I}_n)\times\mathbf{P}_{\mathbf{c}},$$

а в общем случае:

$$\mathbf{P}_{\mathbf{c}}^{\langle -k\rangle}=\mathbf{J}_n\times\mathbf{P}_{\mathbf{q}}^{\langle -k\rangle}=\mathbf{J}_n\times(\mathbf{A}^{\langle -k\rangle}\times\mathbf{P}_{\mathbf{q}})=\mathbf{J}_n\times(\mathbf{A}^{\langle -k\rangle}\times(\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}}))=(\mathbf{J}_n\times\mathbf{A}^{\langle -k\rangle}\times\mathbf{I}_n)\times\mathbf{P}_{\mathbf{c}}.$$

Заметим, подобная замена обратима, более того, $\left(\mathbf{P}_{\mathbf{c}}^{\langle -k\rangle}\right)^{\langle -k\rangle}=\mathbf{P}_{\mathbf{c}}$.

Теперь соединим два полученных выше результата. Пусть поступило свидетельство $\langle m,k\rangle$. Мы считаем что $m\wedge k=0$, то есть в положительную и отрицательную часть свидетельства входят разные атомы. Проведем замену переменных $\langle -k\rangle$, тогда в новых переменных нам надо пропагировать свидетельство $\langle m\vee k\rangle$. Но это мы уже умеем делать. В полученном результате пропагации останется только снова провести замену переменных $\langle -k\rangle$. Итого получаем:

$$\mathbf{P}_{\mathbf{c}}^{\langle m,k\rangle}=\left(\left(\mathbf{P}_{\mathbf{c}}^{\langle m,k\rangle}\right)^{\langle -k\rangle}\right)^{\langle -k\rangle}=\left(\left(\mathbf{P}_{\mathbf{c}}^{\langle -k\rangle}\right)^{\langle m\vee k,0\rangle}\right)^{\langle -k\rangle}=\mathbf{J}_n\times\mathbf{A}^{\langle -k\rangle}\times\mathbf{I}_n\times\left(\mathbf{P}_{\mathbf{c}}^{\langle -k\rangle}\right)^{\langle m\vee k,0\rangle}=$$

$$=\mathbf{J}_n\times\mathbf{A}^{\langle -k\rangle}\times\mathbf{I}_n\times\mathbf{J}_n\times\frac{1}{(\mathbf{1},\mathbf{Q}^{\langle m\vee k,0\rangle}\times\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}}^{\langle -k\rangle})}\cdot\mathbf{Q}^{\langle m\vee k,0\rangle}\times\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}}^{\langle -k\rangle}=$$

$$=\mathbf{J}_n\times\mathbf{A}^{\langle -k\rangle}\times\frac{1}{(\mathbf{1},\mathbf{Q}^{\langle m\vee k,0\rangle}\times\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}}^{\langle -k\rangle})}\cdot\mathbf{Q}^{\langle m\vee k,0\rangle}\times\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}}^{\langle -k\rangle}=$$

$$=\mathbf{J}_n\times\mathbf{A}^{\langle -k\rangle}\times\frac{1}{(\mathbf{1},\mathbf{Q}^{\langle m\vee k,0\rangle}\times\mathbf{I}_n\times\mathbf{J}_n\times\mathbf{A}^{\langle -k\rangle}\times\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}})}\cdot\mathbf{Q}^{\langle m\vee k,0\rangle}\times\mathbf{I}_n\times\mathbf{J}_n\times\mathbf{A}^{\langle -k\rangle}\times\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}}=$$

$$=\mathbf{J}_n\times\mathbf{A}^{\langle -k\rangle}\times\frac{1}{(\mathbf{1},\mathbf{Q}^{\langle m\vee k,0\rangle}\times\mathbf{A}^{\langle -k\rangle}\times\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}})}\cdot\mathbf{Q}^{\langle m\vee k,0\rangle}\times\mathbf{A}^{\langle -k\rangle}\times\mathbf{I}_n\times\mathbf{P}_{\mathbf{c}}.$$

Полученная формула дает нам точное выражение для апостериорной вероятности элементов ФЗ при поступившем свидетельстве, если исходные оценки были стохастические (то есть оценки вероятности конъюнктов имели вид $\mathbf{P}_{\mathbf{c}}[i]=p_i$, где p_i — числа из интервала $[0;1]$). Если оценки были не стохастические, а интервальные, то нам потребуется решать серию ЗЛП [4].

Пусть задан ФЗ с оценками $\mathbf{P}^{-}$ и $\mathbf{P}^{+}$, тогда опишем ЗЛП для получения апостериорной вероятности на элементах ФЗ.

Полученная формула однородна по $\mathbf{P}_{\mathbf{c}}$, то есть если подставить вместо него $\lambda\cdot\mathbf{P}_{\mathbf{c}}$, где $\lambda>0$, то результат не изменится. Рассмотрим в качестве переменных нашей ЗЛП $\lambda\cdot\mathbf{P}_{\mathbf{c}}[i]$ и обозначим их через $\mathbf{d}[i]$. Очевидно, что $\mathbf{d}[0]=\lambda$. Над новыми переменными можно выписать ограничения вида $\mathbf{d}[i]\le\lambda\mathbf{P}^{+}[i]$ и $\lambda\mathbf{P}^{-}[i]\le\mathbf{d}[i]$, кроме этого, мы должны включить ограничения $\mathbf{I}_n\times\mathbf{d}\ge\mathbf{0}$. Кроме указанных ограничений, добавим ограничения $(\mathbf{1}, \mathbf{Q}^{\langle m \vee k, 0 \rangle} \times \mathbf{A}^{\langle -k \rangle} \times \mathbf{I}_n \times \mathbf{d}) = 1$ и $\lambda \geq 0$. Внеся это ограничение, мы ничего не потеряем, так как всегда можно подобрать такое $\lambda > 0$, что это будет верно, а в силу однородности мы не исключим ни одного возможного значения для элементов $\mathbf{P}_{\mathbf{c}}^{\langle m,k \rangle}$ (подробнее применительно к АБС можно посмотреть в [4], а применительно к общему случаю сведения задачи гиперболического программирования к ЗЛП в [8]). Последнее ограничение позволяет переписать выражение для $\mathbf{P}_{\mathbf{c}}^{\langle m,k \rangle}$, а именно

$$\mathbf{P}_{\mathbf{c}}^{\langle m,k \rangle} = \mathbf{J}_n \times A^{\langle -k \rangle} \times \mathbf{Q}^{\langle m \vee k, 0 \rangle} \times A^{\langle -k \rangle} \times \mathbf{I}_n \times \mathbf{d}.$$

Теперь нам осталось только решить полученные ЗЛП и найти при заданных условиях максимумы и минимумы для элементов $\mathbf{P}_{\mathbf{c}}^{\langle m,k \rangle}[i]$, они и будут оценками апостериорной вероятности.

7. Заключение

Мы рассмотрели все основные виды локального логико-вероятностного вывода в АБС, представив соответствующие формулы для вычислений на матрично-векторном языке. Полученные ЗЛП могут быть без особых трудностей решены на ЭВМ, более того, матричные и векторные структуры, описанные в статье, имеют достаточно простое и эффективное машинное представление в виде массивов. Однако открытым остается вопрос, можно ли упростить предложенные матричные операции и тем самым сократить вычислительную сложность алгоритмов.

Литература

1. *Городецкий В. И.* Байесовский вывод. Препринт №149. Л.: ЛИИАН, 1991. 38 с.
2. *Городецкий В. И.* Алгебраические байесовские сети — новая парадигма экспертных систем // Юбилейный сборник трудов институтов Отделения информатики, вычислительной техники и автоматизации РАН. М.: РАН, 1993. Т. 2. С. 120–141.
3. *Nilsson N. J.* Probabilistic Logic // Artificial Intelligence. Amsterdam: Elsevier Science Publishers B.V., 1986. Vol. 47. P. 71–87.
4. *Тулупьев А. Л., Николенко С. И., Сироткин А. В.* Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.
5. *Беллман Р.* Введение в теорию матриц. М.: Наука, Гл. редакция физико-математической литературы, 1969. 368 с.
6. *Тулупьев А. Л.* Алгебраические байесовские сети: глобальный логико-вероятностный вывод. СПб.: СПбГУ; Анатолия. 2007. 40 с.
7. *Сироткин А. В., Тулупьев А. Л.* Локальный априорный вывод в алгебраических байесовских сетях: комплекс основных алгоритмов // Труды СПИИРАН. Вып. 5. 2007. СПб.: Наука, 2007. С. 100–111.
8. *Гавурин М. К., Малоземов В. Н.* Экстремальные задачи с линейными ограничениями: учебное пособие. Л.: Изд-во Ленингр. ун-та, 1984. 175 с.