



УДК 004.8

МАТРИЧНО-ВЕКТОРНЫЕ АЛГОРИТМЫ ЛОКАЛЬНОГО АПОСТЕРИОРНОГО ВЫВОДА В АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЯХ НАД ПРОПОЗИЦИЯМИ-КВАНТАМИ

А.А. Золотин^a, А.Л. Тулупьев^{a, b}, А.В. Сироткин^b^a Санкт-Петербургский государственный университет, Санкт-Петербург, 198504, Российская Федерация^b СПИИРАН, Санкт-Петербург, 199178, Российская Федерация

Адрес для переписки: alexander.tulupyev@gmail.com

Информация о статье

Поступила в редакцию 16.04.15, принята к печати 21.05.15

doi:10.17586/2226-1494-2015-15-4-676-684

Язык статьи – русский

Ссылка для цитирования: Золотин А.А., Тулупьев А.Л., Сироткин А.В. Матрично-векторные алгоритмы локального апостериорного вывода в алгебраических байесовских сетях над пропозициями-квантами // Научно-технический вестник информационных технологий, механики и оптики. 2015. Т. 15. № 4. С. 676–684.

Аннотация

В теории вероятностных графических моделей апостериорный вывод является одним из видов вывода, на котором базируется обработка фрагментов знаний с неопределенностью. В работе рассматривается задача описания локального апостериорного вывода в алгебраических байесовских сетях над фрагментами знаний с пропозициями-квантами с помощью матрично-векторных уравнений, которые существенно опираются на использование тензорного произведения матриц, степени Кронекера, а также произведения Адамара. Получены матричные уравнения для вычисления векторов апостериорных вероятностей в задаче апостериорного вывода над фрагментами знаний с пропозициями-квантами. Аналогичные уравнения того же типа уже рассматривались в рамках теории алгебраических байесовских сетей, однако они были построены только для случая апостериорного вывода во фрагментах знаний над идеалами конъюнктов. При обобщении и развитии матрично-векторного аппарата в части операций с векторами вероятностей пропозиций-квантов был, в частности, адаптирован ряд ранее полученных результатов, касающихся нормирующих множителей в апостериорном выводе и задания линейного проективного оператора с помощью вектора-селектора. Рассмотрены все три вида поступающих свидетельств – детерминированные, стохастические и неточные свидетельства – в сочетании со скалярными и интервальными оценками вероятности истинности пропозициональных формул во фрагментах знаний. Сформированы задачи линейного программирования, решение которых дает интервальные оценки искомых апостериорных вероятностей в случае неточного свидетельства или интервальных оценок во фрагменте знаний. Существование такого описания задачи апостериорного вывода позволяет пополнить набор типов фрагментов знаний, которыми мы можем оперировать в локальном и глобальном апостериорном выводе, а также упростить реализацию комплекса программ за счет возможности использования при программировании на Java, C++ или C# уже существующих сторонних библиотек, эффективно поддерживающих представление и обработку матриц и векторов.

Ключевые слова

байесовские сети, апостериорный вывод, алгоритмы вывода, пропаганда свидетельства, фрагмент знаний над квантами.

Благодарности

Часть результатов, представленных в статье, была получена в рамках исследовательского проекта, поддержанного грантом РФФИ № 15-01-09001-а.

MATRIX-VECTOR ALGORITHMS OF LOCAL POSTERIORI INFERENCE IN ALGEBRAIC BAYESIAN NETWORKS ON QUANTA PROPOSITIONS

A.A. Zolotin^a, A.L. Tulupyev^{a, b}, A.V. Sirotkin^b^a Saint Petersburg State University, Saint Petersburg, 198504, Russian Federation^b SPIIRAS, Saint Petersburg, 199178, Russian Federation

Corresponding author: Alexander.tulupyev@gmail.com

Article info

Received 16.04.15, accepted 21.05.15

doi:10.17586/2226-1494-2015-15-4-676-684

Article in Russian

For citation: Zolotin A.A., Tulupyev A.L., Sirotkin A.V. Matrix-vector algorithms of local posteriori inference in algebraic bayesian networks on quanta propositions. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2015, vol.15, no. 4, pp. 676–684.

Abstract

Posteriori inference is one of the three kinds of probabilistic-logic inferences in the probabilistic graphical models theory and the base for processing of knowledge patterns with probabilistic uncertainty using Bayesian networks. The paper deals with a task of local posteriori inference description in algebraic Bayesian networks that represent a class of probabilistic graphical models by means of matrix-vector equations. The latter are essentially based on the use of tensor product of matrices, Kronecker degree and Hadamard product. Matrix equations for calculating posteriori probabilities vectors within posteriori inference in knowledge patterns with quanta propositions are obtained. Similar equations of the same type have already been discussed within the confines of the theory of algebraic Bayesian networks, but they were built only for the case of posteriori inference in the knowledge patterns on the ideals of conjuncts. During synthesis and development of matrix-vector equations on quanta propositions probability vectors, a number of earlier results concerning normalizing factors in posteriori inference and assignment of linear projective operator with a selector vector was adapted. We consider all three types of incoming evidences - deterministic, stochastic and inaccurate - combined with scalar and interval estimation of probability truth of propositional formulas in the knowledge patterns. Linear programming problems are formed. Their solution gives the desired interval values of posterior probabilities in the case of inaccurate evidence or interval estimates in a knowledge pattern. That sort of description of a posteriori inference gives the possibility to extend the set of knowledge pattern types that we can use in the local and global posteriori inference, as well as simplify complex software implementation by use of existing third-party libraries, effectively supporting submission and processing of matrices and vectors when programming in Java, C++ or C#.

Keywords

Bayesian networks, posteriori inference, inference algorithms, evidence propagation, knowledge pattern over quanta-propositions.

Acknowledgements

The research was partially supported by RFBR Grant No 15-01-09001-a.

Введение

Одним из активно развивающихся направлений информатики и информационных технологий является теория вероятностных графических моделей и разработка программного обеспечения, реализующего полученные результаты этой теории [1]. Байесовские сети, в том числе алгебраические байесовские сети (АБС) [2, 3] и связанные с ними конструкции, являются одним из видов вероятностных графических моделей. Байесовские сети все шире используются в системах распознавания образов [4, 5], системах моделирования поведения [6], системах моделирования рисков и обнаружения отказов [7–9], оценивания повреждений и мониторинга состояния здоровья [10–13], а также при анализе климата и окружающей среды [14, 15].

Успех в приложениях одних видов вероятностных графических моделей, стремление распространить накопленные достижения на решение новых задач, анализ возникающих затруднений, потребность в моделировании различных видов неопределенности данных и знаний приводят к появлению новых теоретических и алгоритмических задач, а также к рассмотрению новых (либо известных, но недостаточно исследованных) моделей указанного класса.

Так, в рамках теории алгебраических байесовских сетей в локальном апостериорном выводе (выводе на основе поступившего свидетельства), а также при попытках распространить влияние свидетельства на соседние фрагменты знаний (пропагировать свидетельство) зачастую возникает необходимость манипулировать не только традиционно рассматриваемыми фрагментами знаний над пропозициями-конъюнктами (подходы описаны в [2, 16–18], матрично-векторные уравнения вывода получены там же), но и над фрагментами знаний с пропозициями-квантами. Оба вида фрагментов знаний обсуждались, например, в [19], но все возникающие задачи локального вывода оценок истинности, в том числе и задача локального апостериорного вывода, во фрагменте знаний с пропозициями-квантами (фрагменте знаний, построенном над пропозициями-квантами) были оставлены практически без рассмотрения, не говоря уже о вопросах описания алгоритмов решения таких задач.

Задача локального апостериорного вывода состоит в том, чтобы автоматически вычислить апостериорные вероятности всех элементов (будь то пропозиции-конъюнкты или пропозиции-кванты) фрагмента знаний при поступившем свидетельстве. С формальной точки зрения фрагмент знаний над квантами родственен традиционно рассматриваемому в теории АБС фрагменту знаний над идеалом конъюнктов с оценками вероятностей их истинности, однако требует иной системы накладываемых на оценки вероятности ограничений, что не позволяет непосредственно воспользоваться формулами апостериорного вывода, полученными ранее [2, 16, 17]. Таким образом, цель настоящей работы состоит в том, чтобы свести вывод апостериорных оценок во фрагменте знаний с пропозициями-квантами к матрично-векторным уравнениям и, в случае интервальных оценок, к использующим эти уравнения экстремальным задачам, на основе преобразования или адаптации уже известных матрично-векторных уравнений локального апостериорного вывода для фрагмента знаний АБС.

Говоря о задаче локального апостериорного вывода, мы, на самом деле, подразумеваем две подзадачи [20]. Первой подзадачей локального апостериорного вывода является оценка вероятности появления свидетельства или кортежа свидетельств над заданным фрагментом знаний. Второй подзадачей апостериорного вывода является оценка апостериорных вероятностей истинности элементов фрагмента знаний

при предположении, что во фрагмент знаний поступило свидетельство. Стоит отметить, что результатами первой подзадачи апостериорного вывода мы воспользуемся при решении второй.

Определения, обозначения и базовые результаты

Чтобы избежать повтора ранее опубликованного материала в заметных объемах, отметим, что в работе мы будем следовать системе терминов и обозначений, сложившихся в [2, 3, 18, 21], а в настоящем разделе постараемся ограничиться введением лишь ключевых для достижения поставленной цели объектов.

Зафиксируем конечное множество атомарных пропозициональных формул (атомов) – упорядоченный алфавит атомов $A = \{v_i\}_{i=0}^{n-1}$. С содержательной точки зрения алфавит атомов представляет набор элементарных высказываний экспертов о предметной области. Определим два набора пропозиций над указанным алфавитом. Первый набор – идеал конъюнктов $\{v_{i_1} v_{i_2} \dots v_{i_k} \mid 0 \leq i_1 < i_2 < \dots < i_k \leq n - 1, 0 \leq k \leq n\}$, где $v_{i_1} v_{i_2} \dots v_{i_k}$ означает конъюнкцию соответствующих атомов, а знак конъюнкции здесь и далее опускается. Таким образом, идеал конъюнктов содержит все возможные конъюнкции атомов, в том числе и вырожденные: пустую конъюнкцию и сами атомы. Литерал \tilde{v}_i означает, что на его месте в пропозициональной формуле может стоять либо v_i , либо его отрицание \bar{v}_i . Тогда, наконец, определим второй набор – множество квантов $Q = \{\tilde{v}_0 \tilde{v}_1 \dots \tilde{v}_{n-1}\}$ над алфавитом $V = \{v_i\}_{i=0}^{n-1}$. Иными словами, квант – это конъюнкция, которая для любого атома из алфавита содержит либо его самого, либо его отрицание. С учетом заданного алфавита каждому из квантов вида $\tilde{v}_0 \tilde{v}_1 \dots \tilde{v}_{n-1}$ можно сопоставить характеристический вектор, а этому вектору, в свою очередь, двоичное число, которое рассматривается как номер кванта (индексирует квант) [2, 3]. Следует отметить, что рассматриваются именно пропозициональные формулы, а не их истинностные значения. Более того, в дальнейших выкладках рассматриваются классы эквивалентности пропозициональных формул; в частности, конъюнкции над одними и теми же формулами, но с разным порядком следования последних, не различаются [2, 3, 8, 12].

Векторы $\mathbf{P}_c = \begin{pmatrix} 1 \\ p(c_1) \\ \vdots \\ p(c_{2^n-1}) \end{pmatrix}$ и $\mathbf{P}_q = \begin{pmatrix} p(q_0) \\ p(q_1) \\ \vdots \\ p(q_{2^n-1}) \end{pmatrix}$ содержат вероятности конъюнктов и квантов соот-

ветственно. Например, для фрагмента знаний, построенного над двумя атомами $\{x_1, x_2\} \subset A$, векторы вероятностей квантов и конъюнктов будут выглядеть следующим образом: $\mathbf{P}_c = \begin{pmatrix} 1 \\ p(x_1) \\ p(x_2) \\ p(x_2 x_1) \end{pmatrix}$ и $\mathbf{P}_q =$

$\begin{pmatrix} p(\bar{x}_2 \bar{x}_1) \\ p(\bar{x}_2 x_1) \\ p(x_2 \bar{x}_1) \\ p(x_2 x_1) \end{pmatrix}$. Эти векторы выражаются друг через друга с помощью соотношений $\mathbf{P}_q = \mathbf{I}_n \mathbf{P}_c$, $\mathbf{P}_c = \mathbf{J}_n \mathbf{P}_q$,

где $\mathbf{I}_n = \mathbf{I}_1^{[n]}$, $\mathbf{I}_1 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, $\mathbf{J}_n = \mathbf{J}_1^{[n]}$, $\mathbf{J}_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ согласно [2, 3, 22]. Показатель $[n]$ используется для обозначения степени Кронекера матриц, а ниже символ \otimes – для обозначения тензорного (прямого) произведения матриц [23]. В случае, когда вероятностные оценки заданы над квантами, мы будем работать либо с фрагментом знаний со скалярными оценками $\langle C, \mathbf{P}_q \rangle$, либо с фрагментом знаний с интервальными оценками $\langle C, \mathbf{P}_q^-, \mathbf{P}_q^+ \rangle$, где C – это носитель фрагмента знаний, т.е. множество квантов, упорядоченных согласно принятой индексации. В случае фрагмента знаний с интервальными оценками допустимое значение $\mathbf{P}_q[i]$ лежит в границах $[\mathbf{P}_q^-[i]; \mathbf{P}_q^+[i]]$. С помощью $[i]$ рядом с вектором мы указываем на i -ю его компоненту. Следует обратить внимание, что индексация компонент начинается с 0.

Под свидетельством мы содержательно понимаем новые «обусловливающие» данные, которые поступили во фрагмент знаний и с учетом которых нам требуется пересмотреть все (или некоторые) оценки [3, 22]. Далее мы с разной степенью детализации рассмотрим все три вида поступающих свидетельств: детерминированное, стохастическое и интервальное [22], вместе с формализацией их вида и содержания.

Фрагмент знаний со скалярными оценками и детерминированное свидетельство

Начнем рассмотрение с пропагации детерминированного свидетельства во фрагменте знаний со скалярными оценками и с носителем C .

Мы говорим, что на вход поступило детерминированное свидетельство, если новые сведения представимы в виде конъюнкции атомов и их отрицаний.

Детерминированное свидетельство можно разбить на положительный и отрицательный конъюнкты и представить в виде пары чисел, в которой одно число указывает на отрицательно означенные, а второе – на положительно означенные атомарные переменные [2, 16]. Как пример рассмотрим свидетельство

$\langle \bar{v}_2 \bar{v}_1 v_0 \rangle = \langle v_0, \bar{v}_2 \bar{v}_1 \rangle = \langle 001_2, 110_2 \rangle = \langle 1, 6 \rangle$. Таким образом, допустим, что во фрагмент знаний поступило детерминированное свидетельство $\langle i; j \rangle$ (иначе можно обозначить как $\langle c_i; c_j \rangle$). Ранее [21] уже были получены и обоснованы результаты для уравнений апостериорного вывода в случае пропагации детерминированного свидетельства во фрагмент знаний со скалярными оценками, поэтому просто приведем здесь соответствующие формулы. Для этого нам понадобятся следующие матрицы:

$$\mathbf{H}^{(i,j)} = \tilde{\mathbf{H}}_{n-1}^{(i,j)} \otimes \tilde{\mathbf{H}}_{n-2}^{(i,j)} \otimes \dots \otimes \tilde{\mathbf{H}}_0^{(i,j)}, \text{ где}$$

$$\tilde{\mathbf{H}}_k^{(i,j)} = \begin{cases} \mathbf{H}^+, & \text{если } v_k \text{ входит в } c_i, \\ \mathbf{H}^-, & \text{если } v_k \text{ входит в } c_j, \\ \mathbf{H}^\circ, & \text{иначе;} \end{cases}$$

причем $\mathbf{H}^+ = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{H}^- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathbf{H}^\circ = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ согласно [16]. Также введем [21] вектор-селектор $\mathbf{s}^{(i,j)}$. Данный вектор интересен тем, что его можно интерпретировать как характеристический вектор, позволяющий отобрать те кванты, которые не противоречат данному свидетельству. С точки зрения программной реализации интересно то, что любой компонент вектора $\mathbf{s}^{(i,j)}$ может быть выражен напрямую через индексы свидетельства $\langle i; j \rangle$. Воспользовавшись вектором $\mathbf{s}^{(i,j)}$, запишем решение первой задачи апостериорного вывода: $p(\langle c_i; c_j \rangle) = (\mathbf{s}^{(i,j)}, \mathbf{P}_q)$, где $\mathbf{s}^{(i,j)} = \otimes_{k=0}^{n-1} \tilde{\mathbf{s}}_k^{(i,j)}$, где

$$\tilde{\mathbf{s}}_k^{(i,j)} = \begin{cases} \mathbf{s}^+, & \text{если } v_k \text{ входит в } c_i, \\ \mathbf{s}^-, & \text{если } v_k \text{ входит в } c_j, \\ \mathbf{s}^\circ, & \text{иначе;} \end{cases}$$

причем $\mathbf{s}^+ = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\mathbf{s}^- = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{s}^\circ = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Здесь и далее запись двух векторов в круглых скобках через запятую обозначает скалярное произведение.

Обозначим искомые апостериорные вероятности квантов за $\mathbf{P}_q^{(i,j)}$. Тогда [2, 21] решение второй задачи апостериорного вывода сводится, в конце концов, к расчету условных вероятностей, при этом вычисления, как и требовалось, при использовании уже введенных обозначений и построенных объектов сведутся к матрично-векторному уравнению

$$\mathbf{P}_q^{(i,j)} = \frac{1}{(\mathbf{s}^{(i,j)}, \mathbf{P}_q)} \cdot \mathbf{H}^{(i,j)} \mathbf{P}_q,$$

где $(\mathbf{s}^{(i,j)}, \mathbf{P}_q)$ – скалярное произведение векторов.

Фактически матрица $\mathbf{H}^{(i,j)}$ задает линейный проективный оператор, который приравнивает нулю вероятности тех квантов, которые логически несовместны с поступившим свидетельством, т.е. конъюнкция каждого из таких квантов и свидетельства является тождественной ложью. Воспользовавшись произведением Адамара (операция покомпонентного произведения векторов одинаковой размерности), этот же линейный проективный оператор можно задать и с использованием вектора-селектора $\mathbf{s}^{(i,j)}$. Для наглядности рассмотрим пример построения матрицы и вектора в случае пропагации свидетельства из одного атома, означенного отрицательно ($-\bar{v}_0$), во фрагмент знаний над множеством атомов $\{v_0, v_1\}$. Тогда

$$\mathbf{s}^{(0,1)} = \mathbf{s}^- \otimes \mathbf{s}^\circ = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \text{ а } \mathbf{H}^{(0,1)} = \mathbf{H}^- \otimes \mathbf{H}^\circ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \text{ Можно заметить, что главная диагональ}$$

матрицы $\mathbf{H}^{(0,1)}$ совпадает с вектором $\mathbf{s}^{(0,1)}$, т.е. $\text{diag} \mathbf{H}^{(0,1)} = \mathbf{s}^{(0,1)}$, что следует из построения каждого из них, поэтому матрично-векторное уравнение можно переписать следующим образом:

$$\mathbf{P}_q^{(i,j)} = \frac{1}{(\mathbf{s}^{(i,j)}, \mathbf{P}_q)} \mathbf{s}^{(i,j)} \circ \mathbf{P}_q, \text{ где символ } \circ \text{ использован для обозначения произведения Адамара. С техно-$$

логической точки зрения применение вектора-селектора облегчает формализацию задачи в системах, ориентированных на работу с векторами и матрицами, например, в системе MATLAB и системе R.

Экстремальные задачи в локальном апостериорном выводе

Итак, перед нами стоит следующая задача: выразить на матрично-векторном языке вероятность поступившего свидетельства $p(\langle c_i; c_j \rangle)$ для различных видов свидетельств, а также составить и обосновать задачу линейного программирования, решением которой будут интервальные оценки апостериорных вероятностей $\mathbf{P}_q^{(i,j)}$ для фрагмента знаний над квантами. Воспользовавшись результатами, полученными для фрагментов знаний над идеалом конъюнктов [17, 18], мы хотим решить поставленные задачи по аналогии с [2, 16].

В настоящем разделе мы рассматриваем задачу, когда оценки вероятностей во фрагменте знаний с квантами – интервальные, т.е. задан непротиворечивый фрагмент знаний с интервальными оценками вероятностей $\langle C, \mathbf{P}_q^-, \mathbf{P}_q^+ \rangle$. По аналогии с предыдущим случаем предположим, что поступило детерминиро-

ванное свидетельство $\langle i; j \rangle$. Как и в случае с фрагментом знаний со скалярными оценками, вероятность поступившего свидетельства можно вычислить по следующей формуле: $p(\langle c_i; c_j \rangle) = (\mathbf{s}^{(i,j)}, \mathbf{P}_q)$. Однако вследствие того, что значение каждой из компонент вектора \mathbf{P}_q лежит в интервале $[\mathbf{P}_q^-[i], \mathbf{P}_q^+[i]]$, решение первой задачи апостериорного вывода в случае детерминированного свидетельства и фрагмента знаний с интервальными оценками сводится к решению задач линейного программирования по нахождению максимума и минимума значения $(\mathbf{s}^{(i,j)}, \mathbf{P}_q)$. Стоит также отметить, что на вектор переменных \mathbf{P}_q накладываются линейные ограничения для поддержания непротиворечивости фрагмента знаний: ограничения, вытекающие из аксиоматики теории вероятностей (\mathcal{E}), и ограничения, вытекающие из предметной области (\mathcal{D}). Ограничение \mathcal{E} в случае фрагмента знаний над квантами состоит из двух условий, записываемых на матрично-векторном языке: $\{\mathbf{P}_q \geq \mathbf{0}, (\mathbf{1}, \mathbf{P}_q) = 1\}$, а ограничение \mathcal{D} может быть записано как совокупность условий: $\forall i \in [0..n-1] \mathbf{P}_q^-[i] \leq \mathbf{P}_q[i] \leq \mathbf{P}_q^+[i]$, либо, как это принято в математических дисциплинах, посвященных экстремальным задачам, $\mathbf{P}_q^- \leq \mathbf{P}_q \leq \mathbf{P}_q^+$, где n – это мощность множества квантов, над которым построен данный фрагмент знаний, т.е. множество ограничений $\mathcal{D} = \{\mathbf{P}_q^- \leq \mathbf{P}_q \leq \mathbf{P}_q^+\}$. Таким образом, решением первой задачи апостериорного вывода является замкнутый промежуток, ограниченный следующими значениями:

$$\left(\min_{\mathcal{D} \cup \mathcal{E}} (\mathbf{s}^{(i,j)}, \mathbf{P}_q) \right) \text{ и } \left(\max_{\mathcal{D} \cup \mathcal{E}} (\mathbf{s}^{(i,j)}, \mathbf{P}_q) \right).$$

Указанные минимум и максимум находятся как результаты решения задач линейного программирования.

Так как во фрагменте знаний с квантами нам заданы интервальные оценки вероятности их истинности, мы имеем дело с семейством распределений вероятностей; следовательно, искомые во второй подзадаче апостериорные вероятности могут быть оценены лишь интервально. Поиск верхней и нижней границы $\mathbf{P}_q^{(i,j)}$ сводится к поиску минимума и максимума выражения $\mathbf{P}_q^{(i,j)} = \frac{1}{(\mathbf{s}^{(i,j)}, \mathbf{P}_q)} \cdot \mathbf{s}^{(i,j)} \circ \mathbf{P}_q$ при ограничениях $\mathcal{D} \cup \mathcal{E}$. Стоит отметить, что здесь и далее задача о поиске минимума и максимума решается независимо и отдельно для каждой целевой функции – компоненты вектора $\mathbf{P}_q^{(i,j)}$ – относительно переменных вектора \mathbf{P}_q .

Экстремальные задачи по нахождению минимума и максимума указанных целевых функций относятся к известному классу задач дробно-линейного программирования. Для сведения ее к задаче линейного программирования нам нужно избавиться от множителя $\frac{1}{(\mathbf{s}^{(i,j)}, \mathbf{P}_q)}$. Воспользовавшись тем же приемом и обоснованием, что и в [2, 3, 16], зададим новый вектор переменных (и, таким образом, произведем замену переменных) $\mathbf{D} = \lambda \mathbf{P}_q$, тогда $\frac{1}{(\mathbf{s}^{(i,j)}, \mathbf{P}_q)} \cdot \mathbf{s}^{(i,j)} \circ \mathbf{P}_q = \frac{1}{(\mathbf{s}^{(i,j)}, \lambda \mathbf{P}_q)} \cdot \mathbf{s}^{(i,j)} \circ \lambda \mathbf{P}_q = \frac{1}{(\mathbf{s}^{(i,j)}, \mathbf{D})} \cdot \mathbf{s}^{(i,j)} \circ \mathbf{D}$. Теперь положим $\frac{1}{(\mathbf{s}^{(i,j)}, \mathbf{D})} = 1$. При новых переменных и введенных ограничениях исходные экстремальные задачи сведутся к задачам линейного программирования – поиску минимума и максимума компонент вектора $\mathbf{s}^{(i,j)} \circ \mathbf{D}$ при условиях $\mathcal{D}' \cup \mathcal{E}' \cup \{(\mathbf{s}^{(i,j)}, \mathbf{D}) = 1\}$.

Причем $\mathcal{D}' = \{\lambda \mathbf{P}_q^- \leq \mathbf{D}\} \cup \{\lambda \mathbf{P}_q^+ \geq \mathbf{D}\}$, $\mathcal{E}' = \{\mathbf{D} \geq \mathbf{0}\} \cup \{\lambda \geq 0\} \cup \{(\mathbf{D}, \mathbf{1}) = \lambda\}$. Как и в решении первой задачи апостериорного вывода, данная совокупность линейных ограничений обеспечивает непротиворечивость фрагмента знаний над пропозициями-квантами с новыми, т.е. апостериорными, оценками вероятностей. Здесь, аналогично предыдущему случаю, в качестве переменных выступают компоненты вектора \mathbf{D} и λ .

Перед рассмотрением следующего случая дадим определение второго вида свидетельств – стохастического.

Изначально, в рамках теории АБС, стохастическое свидетельство – это предположение о том, что на подыдеале C' носителя фрагмента знаний C задан непротиворечивый фрагмент знаний со скалярными оценками, который определяет вероятности истинности элементов соответствующего подыдеала. По аналогии, рассматривая стохастическое свидетельство над пропозициями-квантами, вероятность пропозиций-квантов свидетельства обозначим как \mathbf{P}_q^a .

Теперь предположим, что во фрагмент знаний со скалярными оценками поступило стохастическое свидетельство с вероятностями пропозиций-квантов \mathbf{P}_q^a . Рассмотрение пропагации стохастического свидетельства сведем к рассмотрению пропагации серии детерминированных свидетельств, на которые его можно «разбить» [22]. Если вероятность какого-то детерминированного свидетельства равна 0, но при этом оценки, задаваемые стохастическим свидетельством для данного детерминированного свидетельства, нулю не равны, то такое свидетельство называется невероятным.

Отметим, что поступившему стохастическому свидетельству можно однозначно сопоставить вектор вероятностей конъюнктов \mathbf{P}_c^a (а ему, в свою очередь, однозначно соответствует вектор \mathbf{P}_q^a , такой, что

$\mathbf{P}_c^a = \mathbf{J}_n \mathbf{P}_q^a$). Пропагация свидетельства, построенного над идеалом конъюнктов, рассматривалась нами ранее и по своей сути сводится к свидетельству над пропозициями-квантами. Это можно легко заметить, приняв во внимание то, что для формирования конечного результата апостериорного вывода мы умножаем результат пропагации каждого из детерминированных свидетельств на соответствующий элемент вектора \mathbf{P}_q^a , а затем складываем полученные произведения и, окончательно, формируем таким образом результат апостериорного вывода [22]. Также для сопоставления индексов детерминированных свидетельств с индексами множества квантов поступившего свидетельства введем функцию $\text{GInd}(i, m)$ [2, 16], где i – индекс кванта в алфавите, над которым построено свидетельство, а m – индекс наибольшего элемента \mathbf{P}_q^a в исходном алфавите. Тогда, обозначив мощность алфавита, над которым построен данный фрагмент знаний, за n' , мы можем составить матрично-векторное уравнение для решения первой задачи апостериорного вывода. Вероятность поступившего свидетельства оценивается следующим значением:

$$\sum_{i=0}^{2^{n'}-1} \left(\mathbf{s}^{\langle \text{GInd}(i,m), \text{GInd}(2^{n'}-1-i,m) \rangle}, \mathbf{P}_q \right) \mathbf{P}_q^a [i].$$

В решении второй задачи апостериорного вывода воспользуемся тем же приемом и рассмотрим стохастическое свидетельство как совокупность детерминированных. Пользуясь формулами [2, 16, 21], полученными при решении второй задачи апостериорного вывода в случае детерминированного свидетельства, решением указанной задачи в рамках фрагмента знаний со скалярными оценками будет следующая сумма:

$$\mathbf{P}_q^{(i;j)} = \sum_{i=0}^{2^{n'}-1} \frac{\mathbf{s}^{\langle \text{GInd}(i,m), \text{GInd}(2^{n'}-1-i,m) \rangle} \circ \mathbf{P}_q}{\left(\mathbf{s}^{\langle \text{GInd}(i,m), \text{GInd}(2^{n'}-1-i,m) \rangle}, \mathbf{P}_q \right)} \cdot \mathbf{P}_q^a [i].$$

В случае интервальных оценок во фрагменте знаний мы аналогичным образом представляем стохастическое свидетельство в качестве серии детерминированных, однако в данном случае решениями первой и второй задач апостериорного вывода будут интервальные накрывающие оценки вероятностей. Они получатся как линейная комбинация интервальных оценок, появляющихся при разложении стохастического свидетельства на детерминированные свидетельства и их последующее пропагирование. Проведя вышеуказанные операции, мы получим следующие уравнения для решения задач апостериорного вывода. Решением первой задачи будет интервал, ограниченный данными значениями:

$$\min_{\mathcal{D} \cup \mathcal{E}} \sum_{i=0}^{2^{n'}-1} \left(\mathbf{s}^{\langle \text{GInd}(i,m), \text{GInd}(2^{n'}-1-i,m) \rangle}, \mathbf{P}_q \right) \mathbf{P}_q^a [i] \text{ и}$$

$$\max_{\mathcal{D} \cup \mathcal{E}} \sum_{i=0}^{2^{n'}-1} \left(\mathbf{s}^{\langle \text{GInd}(i,m), \text{GInd}(2^{n'}-1-i,m) \rangle}, \mathbf{P}_q \right) \mathbf{P}_q^a [i].$$

Для решения второй задачи потребуется решить задачу линейного программирования, покомпонентно найдя минимум и максимум данного выражения для каждого из детерминированных свидетельств, а затем просуммировать по всем индексам.

$$\mathbf{P}_q^{(i;j)} = \mathbf{s}^{\langle \text{GInd}(i,m), \text{GInd}(2^{n'}-1-i,m) \rangle} \circ \mathbf{D}$$

при условии $\mathcal{R}' = \mathcal{D}' \cup \mathcal{E}' \cup \left\{ \left(\mathbf{s}^{\langle \text{GInd}(i,m), \text{GInd}(2^{n'}-1-i,m) \rangle}, \mathbf{D} \right) = 1 \right\}$. Причем $\mathcal{D}' = \{ \lambda \mathbf{P}_q^- \leq \mathbf{D} \} \cup \{ \lambda \mathbf{P}_q^+ \geq \mathbf{D} \}$, а $\mathcal{E}' = \{ \mathbf{D} \geq \mathbf{0} \} \cup \{ \lambda \geq 0 \} \cup \{ (\mathbf{D}, \mathbf{1}) = \lambda \}$.

Таким образом, конечное решение второй задачи апостериорного вывода в случае стохастического свидетельства будет выглядеть следующим образом:

$$\mathbf{P}_q^{(i;j), \min} = \sum_{i=0}^{2^{n'}-1} \min_{\mathcal{R}'} \left(\mathbf{s}^{\langle \text{GInd}(i,m), \text{GInd}(2^{n'}-1-i,m) \rangle} \circ \mathbf{D} \right) \mathbf{P}_q^a [i],$$

$$\mathbf{P}_q^{(i;j), \max} = \sum_{i=0}^{2^{n'}-1} \max_{\mathcal{R}'} \left(\mathbf{s}^{\langle \text{GInd}(i,m), \text{GInd}(2^{n'}-1-i,m) \rangle} \circ \mathbf{D} \right) \mathbf{P}_q^a [i].$$

Наконец, рассмотрим третий вид свидетельства – интервальное (неточное).

Неточное свидетельство представляет собой фрагмент знаний над квантами с интервальными оценками $\langle C, \mathbf{P}_q^{a,-}, \mathbf{P}_q^{a,+} \rangle$. Содержательно пропагация интервального свидетельства во фрагменты знаний со скалярными и интервальными оценками сводится к решению задачи линейного программирования для поиска интервальных оценок. Обработка интервального свидетельства в случае фрагмента знаний над идеалом конъюнктов описана в [2, 16], при этом обработка пропагации интервальных свидетельств во фрагмент знаний над квантами будет аналогична. К сожалению, при решении первой задачи апостериорного вывода в случае неточных свидетельств и фрагмента знаний с интервальными оценками (как и в случае стохастического свидетельства) мы сталкиваемся с выходом за пределы задач линейного программирования; решение же последних приводит лишь к накрывающим оценкам [2, 16].

Пример

В предыдущих разделах нами был подробно рассмотрен локальный апостериорный вывод для фрагмента знаний над пропозициями-квантами с интервальными оценками вероятностей и даны форму-

лы для решения обеих задач апостериорного вывода для различных видов поступающих свидетельств. Ниже будут предложены примеры использования описанного ранее вектора-селектора и проведено сравнение объемов вычислений с и без его использования.

Для большей наглядности рассмотрим пропагацию детерминированного свидетельства, построенного над двумя атомами $A = \{v_0, v_2\}$, во фрагмент знаний $\langle C, P_q \rangle$ со скалярными оценками вероятности истинности пропозиций-квантов, построенный над тремя атомами $A = \{v_0, v_1, v_2\}$. Пусть вероятности оценок истинности множества пропозиций-квантов являются непротиворечивыми и равны следующим значениям:

$$P_q = \begin{pmatrix} 0,3 \\ 0,12 \\ 0,07 \\ 0,09 \\ 0,15 \\ 0,01 \\ 0,1 \\ 0,15 \end{pmatrix}, \text{ а поступившее свидетельство записывается как } \langle v_2 \bar{v}_0 \rangle = \langle 100,001 \rangle = \langle 4,1 \rangle. \text{ Вычислим век-}$$

$$\text{тор-селектор } \mathbf{s}^{(4;1)} = \mathbf{s}^- \otimes \mathbf{s}^o \otimes \mathbf{s}^+ = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \text{ и соответствующую данному свидетельству диагональную}$$

$$\text{матрицу } \mathbf{H}^{(4,1)} = \mathbf{H}^- \otimes \mathbf{H}^o \otimes \mathbf{H}^+ = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Тогда $p\langle 4,1 \rangle = (\mathbf{1}, \mathbf{H}^{(4,1)} P_q) = (\mathbf{s}^{(4,1)}, P_q) = 0,25$, а решением второй задачи апостериорного вывода

$$\text{будет вектор } P_q^{(4,1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0,15 \\ 0 \\ 0,1 \\ 0 \end{pmatrix}. \text{ Очевидно, что конечные результаты при использовании вектора-селектора}$$

и без его использования эквивалентны, однако при использовании вектора-селектора не приходится прибегать к особым приемам работы с сильно разреженными матрицами: вектор содержит все необходимые сведения. Говоря об объеме требуемой памяти и вычислений, можно привести пример из родственных АБС байесовских сетей доверия: пусть имеется сеть локаторов с 1 миллионом узлов, тогда свидетельства, поступающие с локаторов, могут достигать размера в 100 000 элементов. Таким образом, используя вектор-селектор вместо матрицы $100\ 000 \times 100\ 000$, нам требуется вычислить лишь вектор $1 \times 100\ 000$.

Заключение

В настоящей работе был рассмотрен локальный апостериорный логико-вероятностный вывод во фрагментах знаний, построенных над квантами. Алгоритмизация этого вида логико-вероятностного вывода осуществляется с помощью матрично-векторных уравнений, которые существенно опираются на использование тензорного произведения матриц, степени Кронекера, а также произведение Адамара. Наличие такого описания существенно упрощает программную реализацию алгоритмов вывода и подбор структур данных. Более того, эффективность сведения решения задач к матрично-векторным операциям наиболее ярко проявляется при разработке кода в системах R и MATLAB, ориентированных именно на оптимизацию представления и обработки данных в матричной форме. Наконец, матрично-векторное представление дает нам новые подходы к решению задач, связанных с пропагацией свидетельства, а особенность записи всех результатов в матрично-векторной форме упрощает реализацию алгоритмов за счет возможности использования при программировании на Java, C++ или C# уже существующих сторонних библиотек, эффективно поддерживающих представление и обработку матриц и векторов. Следует отме-

тять, что в настоящей работе оставлены без рассмотрения свойства полученных результатов (например, входят или не входят границы в интервальную оценку апостериорных вероятностей; нет ли лагун в самом интервале-оценке; как должно интерпретироваться и обрабатываться невероятное свидетельство и т.д.). Мы полагаем, что в силу того, что, в конце концов, операции ведутся над выпуклыми множествами, эти свойства будут совпадать с соответствующими им свойствами алгоритмов и результатов локального апостериорного вывода в случае традиционного фрагмента знаний над конъюнктами [2, 3, 22], а строгое полномасштабное обоснование своей точки зрения не приводим, стремясь ограничить объем изложения и сохранить концентрацию на заданной цели, при этом оставляя само обоснование для изложения в последующих работах.

References

1. Arsene O., Dumitrache L., Miha I. Expert system for medicine diagnosis using software agents. *Expert Systems with Applications*, 2015, vol. 42, no. 4, pp. 1825–1834. doi: 10.1016/j.eswa.2014.10.026
2. Tulup'ev A.L., Sirotkin A.V., Nikolenko S.I. *Baiesovskie Seti Doveriya: Logiko-Veroyatnostnyi Vывod v Atsiklicheskih Napravlennykh Grafakh* [Bayesian Belief Networks: Logical-Probabilistic Inference in the Acyclic Directed Graph]. St. Petersburg, SPbSU Publ., 2009, 400 p.
3. Tulup'ev A.L., Nikolenko S.I., Sirotkin A.V. *Baiesovskie Seti: Logiko-Veroyatnostnyi Podkhod* [Bayesian Networks: Logical-Probabilistic Approach]. St. Petersburg, Nauka Publ., 2006, 607 p.
4. Shang J.D., Wang Z.L., Huang Q. A robust algorithm for joint sparse recovery in presence of impulsive noise. *IEEE Signal Processing Letters*, 2015, vol. 22, no. 8, pp. 1166–1170. doi: 10.1109/LSP.2014.2387435
5. Peng P., Tian Y., Wang Y., Li J., Huang T. Robust multiple cameras pedestrian detection with multi-view Bayesian network. *Pattern Recognition*, 2015, vol. 48, no. 5, pp. 1760–1772. doi: 10.1016/j.patcog.2014.12.004
6. Dabrowski J.J., Villiers J.P. Maritime piracy situation modelling with dynamic Bayesian networks. *Information Fusion*, 2015, vol. 23, pp. 116–130. doi: 10.1016/j.inffus.2014.07.001
7. Leu S.S., Chang C.M. Bayesian-network-based fall risk evaluation of steel construction projects by fault tree transformation. *Journal of Civil Engineering and Management*, 2015, vol. 21, pp. 334–342.
8. Mahboob Q., Schone E., Maschek U., Trinckauf J. Investment into human risks in railways and decision optimization. *Human and Ecological Risk Assessment*, 2015, vol. 21, no. 5, pp. 1299–1313. doi: 10.1080/10807039.2014.958375
9. Shin J., Son H., Khalil Ur R., Heo G. Development of a cyber security risk model using Bayesian networks. *Reliability Engineering System Safety*, 2015, vol. 134, pp. 208–217. doi: 10.1016/j.res.2014.10.006
10. Arangio S., Bontempi F. Structural health monitoring of a cable-stayed bridge with Bayesian neural networks. *Structure and Infrastructure Engineering*, 2015, vol. 11, no. 4, pp. 575–587. doi: 10.1080/15732479.2014.951867
11. Dawson P., Gailis R., Meehan A. Detecting disease outbreaks using a combined Bayesian network and particle filter approach. *Journal of Theoretical Biology*, 2015, vol. 370, pp. 171–183. doi: 10.1016/j.jtbi.2015.01.023
12. Chang Y.-S., Fan C.-T., Lo W.-T., Hung W.-C., Yuan S.-M. Mobile cloud-based depression diagnosis using an ontology and a Bayesian network. *Future Generation Computer Systems*, 2015, vol. 43–44, pp. 87–98. doi: 10.1016/j.future.2014.05.004
13. Rafiq M.I., Chryssanthopoulos M.K., Sathanathan S. Bridge condition modelling and prediction using dynamic Bayesian belief networks. *Structure and Infrastructure Engineering*, 2015, vol. 11, no. 1, pp. 38–50. doi: 10.1080/15732479.2013.879319
14. Hamilton S.H., Pollino C.A., Jakeman A.J. Habitat suitability modelling of rare species using Bayesian networks: model evaluation under limited data. *Ecological Modelling*, 2015, vol. 299, pp. 64–78. doi: 10.1016/j.ecolmodel.2014.12.004
15. Liu K.F.-R., Kuo J.-Y., Yeh K., Chen C.-W., Liang H.-H., Sun Y.-H. Using fuzzy logic to generate conditional probabilities in Bayesian belief networks: a case study of ecological assessment. *International Journal of Environmental Science and Technology*, 2015, vol. 12, no. 3, pp. 871–884. doi: 10.1007/s13762-013-0459-x
16. Sirotkin A.V. *Algebraicheskie Baiesovskie Seti: Vychislitel'naya Slozhnost' Algoritmov Logiko-Veroyatnostnogo Vывoda v Usloviyakh Neopredelennosti. Dis. kand. fiz.-mat. nauk* [Algebraic Bayesian Networks: Computational Complexity of Algorithms of Logic-Probabilistic Inference Under Uncertainty. Diss. Ph.D. in Phys.-Math. Sci.]. St. Petersburg, SPbGU, 2011, 218 p.
17. Tulup'ev A.L. *Algebraicheskie Baiesovskie Seti: Logiko-Veroyatnostnaya Graficheskaya Model' Baz Fragmentov Znaniy s Neopredelennost'yu. Dis. dr. fiz.-mat. nauk* [Algebraic Bayesian Networks: Logical-Probabilistic Graphical Model Database Fragments of Knowledge with Uncertainty. Diss. Dr. in Phys.-Math. Sci.]. St. Petersburg, SPbGU, 2009, 670 p.

18. Tulup'ev A.L., Sirotkin A.V. Matrichnye uravneniya lokal'nogo logiko-veroyatnostnogo vyvoda otsenok istinnosti elementov v algebraicheskikh baiesovskikh setyakh [Matrix-Vector Equations for Local Probabilistic-Logic Inference in Algebraic Bayesian Networks]. *Vestnik of the St. Petersburg University: Mathematics*, 2012, no. 3, pp. 63–72.
19. Tulup'ev A.L. *Baiesovskie Seti: Logiko-Veroyatnostnyi Vyvod v Tsiklakh* [Bayesian Networks: Logic-Probabilistic Inference in Cycles]. St. Petersburg, SPbGU, 2008, 140 p.
20. Tulup'ev A.L. Aposteriornye otsenki veroyatnostei v algebraicheskikh baiesovskikh setyakh [A posteriori probabilistic estimates in algebraic Bayesian networks]. *Vestnik of the St. Petersburg University: Seriya 10: Prikladnaya Matematika. Informatika. Protsessy Upravleniya*, 2012, no. 2, pp. 51–59.
21. Zolotin A.A., Tulupyev A.L., Sirotkin A.V. Matrix-vector algorithms for normalizing factors in algebraic Bayesian networks local posteriori inference. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2015, vol. 15, no. 1, pp. 78–85 (in Russian)
22. Tulup'ev A.L. *Algebraicheskie Baiesovskie Seti: Lokal'nyi Logiko-Veroyatnostnyi Vyvod* [Algebraic Bayesian Networks: a Local Logic-Probabilistic Inference]. St. Petersburg, SPbGU Publ., Anatoliya, 2007, 80 p.
23. Bellman R. *Introduction to Matrix Analysis*. NY, McGraw-Hill, 1970.

- | | |
|--|---|
| <i>Золотин Андрей Александрович</i> | – студент, Санкт-Петербургский государственный университет, Санкт-Петербург, 198504, Российская Федерация, andrey.zolotin@gmail.com |
| <i>Тулупьев Александр Львович</i> | – доктор физико-математических наук, доцент, заведующий лабораторией, СПИИРАН, Санкт-Петербург, 199178, Российская Федерация; профессор, Санкт-Петербургский государственный университет, Санкт-Петербург, 198504, Российская Федерация, alexander.tulupyev@gmail.com |
| <i>Сироткин Александр Владимирович</i> | – кандидат физико-математических наук, старший научный сотрудник, СПИИРАН, Санкт-Петербург, 199178, Российская Федерация, alexander.sirotkin@gmail.com |
| <i>Andrey A. Zolotin</i> | – student, Saint Petersburg State University, Saint Petersburg, 198504, Russian Federation, andrey.zolotin@gmail.com |
| <i>Alexander L. Tulupyev</i> | – D.Sc., Associate professor, Laboratory head, SPIIRAS, Saint Petersburg, 199178, Russian Federation; Professor, Saint Petersburg State University, Saint Petersburg, 198504, Russian Federation, alexander.tulupyev@gmail.com |
| <i>Alexander V. Sirotkin</i> | – PhD, senior scientific researcher, SPIIRAS, Saint Petersburg, 199178, Russian Federation, alexander.sirotkin@gmail.com |