ПРОБЛЕМЫ ФУНДАМЕНТАЛЬНЫХ И ПРИКЛАДНЫХ НАУК

УДК 519.226

Н.В. Кузнєцова, магістр, Нац. техн. ун-т України "КПІ"

ГІБРИДНІ МЕРЕЖІ БАЙЄСА: ОСНОВНІ ОСОБЛИВОСТІ І ТОЧНІ МЕТОДИ ФОРМУВАННЯ ВИСНОВКУ

Н.В. Кузнєцова. **Гібридні мережі Байєса: основні особливості і точні методи формування висновку.** Розглянуто особливості гібридних мереж Байєса (ГМБ) і основні їх види. Проаналізовано основні методи формування точного висновку для мереж Байєса і показана їх адаптація для ГМБ.

Н.В. Кузнецова. **Гибридные сети Байеса: основные особенности и точные методы формирования вывода.** Рассмотрены особенности гибридных сетей Байеса (ГСБ) и основные их виды. Проанализированы методы формирования точного вывода для сетей Байеса и показана их адаптация для ГСБ.

N.V. Kuznyetsova. **Hybrid Bayesian Networks: the main features and exact inference methods.** The features of Hybrid Bayesian networks (HBN) and their basic types are considered. The exact inference methods for Bayesian Networks are analyzed and their adaptation for HBN is shown.

Сьогодні, коли з кожним днем збільшуються обсяги та кількість операцій з даними, потрібно використовувати нові інструменти для спрощення обробки даних. Раніше розглядалися моделі, що включали лише дискретні змінні [1, 2]. Однак існує багато реальних галузей, модельованих як гібридні стохастичні процеси, які включають як неперервні змінні, так і дискретні. Такі процеси наявні у разі відслідковування цілі, у завданнях візуального відслідковування, у задачах швидкісного розпізнання мови та інших.

Тому важливим є описання властивостей гібридних мереж, а також особливостей їх побудови і формування висновку. Необхідно проаналізувати основні алгоритми формування точного висновку для мереж Байєса (МБ), їх адаптацію та специфіку застосування для гібридних мереж.

Гібридна мережа Байєса $\mathrm{B} = (X, D, P)$ визначається через направлений ациклічний граф $G = (X, E)$ і його функції $P_i = \{P(x_i \mid pa_i)\}$, де pa_i — набір батьківських вузлів x_i. X — набір змінних, розділених на дискретні Δ і неперервні Γ змінні, тобто $X = \Gamma \bigcup \Delta$. Структура графа G обмежена тим, що неперервні змінні не можуть мати дискретні змінні, як їх вузли-нащадки [3]. Умовний розподіл неперервних змінних задається лінійною гаусівською моделлю

$$P(x_i \mid I = i, Z = z) = N(\alpha(i) + \beta(i) \times z, \gamma(i)) \quad x_i \in \Gamma,$$

де Z та I — набори відповідно неперервних і дискретних батьків x_i;

$N(\mu, \sigma)$ — мультиваріантний нормальний розподіл.

Мережа являє собою сумісний розподіл усіх його змінних, заданих добутком усіх його таблиць умовних ймовірностей.

Для одновимірного випадку нормальний розподіл характеризують двома параметрами — середнім μ і дисперсією σ^2. Для зручності розподіл параметризують, використовуючи дисперсію, що є квадратом стандартного відхилення. Функція щільності має вигляд

$$P(X) = N(X;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \quad (1)$$

У багатовимірному випадку нормальний розподіл характеризується двома параметрами — вектором середнього і коваріаційною матрицею. Кажуть, що випадкові змінні $X = X_1,...,X_n$ мають нормальний розподіл $N(X;\boldsymbol{\mu},\boldsymbol{\Sigma})$ (де $\boldsymbol{\mu}$ — вектор розміром n; $\boldsymbol{\Sigma}$ — симетрична позитивно визначена матриця розміром $n \times n$), якщо

$$P(X) = N(X;\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2}\,|\boldsymbol{\Sigma}|^{1/2}}\exp\left(-\frac{1}{2}(x-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(x-\boldsymbol{\mu})\right). \quad (2)$$

Контур багатовимірної гаусіани виглядає як еліпсоїди навколо середнього $\boldsymbol{\mu}$, де кожний еліпсоїд має певну сталу щільність. Вектор середнього визначає центр еліпсоїдів і коваріаційна матриця визначає їх обрис. Взагалі, $\sum_{i,i}$ — це відхилення X_i, $\sum_{i,j} = \sum_{j,i}$ (для $i \neq j$) — це коваріація між X_i і X_j. Вектор середнього і коваріаційна матриця відповідають першим двом моментам нормального розподілу, тобто $\boldsymbol{\mu} = E[X]$; $\boldsymbol{\Sigma} = E[XX^T] - E[X]E[X]^T$. Якщо всі недіагональні елементи у матриці нульові, тоді усі змінні незалежні одна від одної, а гаусіана є *сферичною*.

Зазвичай розглядається сумісний нормальний розподіл над $\{X,Y\}$, де $X \in \Re^n$ і $Y \in \Re^n$. Використовується позначення

$$P(X,Y) = N\left(\begin{pmatrix}\mu_X\\ \mu_Y\end{pmatrix},\begin{bmatrix}\boldsymbol{\Sigma}_{XX}\boldsymbol{\Sigma}_{XY}\\ \boldsymbol{\Sigma}_{YX}\boldsymbol{\Sigma}_{YY}\end{bmatrix}\right), \quad (3)$$

де $\mu_X \in \Re^n$, $\mu_Y \in \Re^n$,

$\boldsymbol{\Sigma}_{XX}$ — матриця розміром $n \times n$,

$\boldsymbol{\Sigma}_{XY}$ — матриця розміром $n \times m$,

$\boldsymbol{\Sigma}_{XY} = \boldsymbol{\Sigma}_{YX}^T$ — матриця розміром $n \times m$,

$\boldsymbol{\Sigma}_{YY}$ — матриця розміром $m \times m$.

Нехай $\{X,Y\}$ має сумісний нормальний розподіл, визначений у формулі (3). Умовний розподіл $P(Y \mid X)$ — це нормальний розподіл $N(Y;\mu'_Y,\boldsymbol{\Sigma}_{YY}^{-1})$, де $\mu'_Y = \mu_Y + \boldsymbol{\Sigma}_{YX}\boldsymbol{\Sigma}_{XX}^{-1}(x-\mu_X)$, $\boldsymbol{\Sigma}'_{YY} = \boldsymbol{\Sigma}_{YY} - \boldsymbol{\Sigma}_{YX}\boldsymbol{\Sigma}_{XX}^{-1}\boldsymbol{\Sigma}_{XY}$.

Для окремого випадку, коли $|Y| = 1$, справедливий висновок: умовний розподіл $P(Y \mid X)$ — це нормальний розподіл $N(Y;\boldsymbol{\beta}_0 + \boldsymbol{\beta}^{\mathrm{T}}X,\sigma^2)$,

де $\boldsymbol{\beta}_0 = \mu_Y - \boldsymbol{\Sigma}_{YX}\boldsymbol{\Sigma}_{XX}^{-1}\mu_X$,

$\boldsymbol{\beta} = \boldsymbol{\Sigma}_{YY} - \boldsymbol{\Sigma}_{YX}\boldsymbol{\Sigma}_{XX}^{-1}\boldsymbol{\Sigma}_{XY}$,

$\boldsymbol{\beta}^{\mathrm{T}}$ — транспонована матриця $\boldsymbol{\beta}$.

Можна розглядати сформульований висновок як шлях конвертації багатовимірних гаусіан у мережу Байєса. Змінні впорядковуються у певному порядку $X_1,...,X_n$. Потім використовується цей висновок для знаходження умовного розподілу

$$P(X_i \mid X_1,...,X_{i-1}) = N\left(X_i;\beta_{i,0} + \sum_{j=1}^{i-1}\beta_{i,j}X_j,\sigma_i^2\right). \quad (4)$$

Створюється ребро з X_j до X_i ($1 \le j < i$) т.т.т.к. $\beta_{i,j} \neq 0$. Умовний ймовірнісний розподіл (УЙР) X_i називається *лінійним умовним ймовірнісним розподілом* і має вигляд виразу (4) (після скорочення усіх нульових $\beta_{i,j}$). Лінійний умовний ймовірнісний розподіл для кореневих вузлів є просто одновимірною гаусіаною. Мережа Байєса, у якій всі умовні ймовірнісні розподіли є лінійними, називається *лінійною гаусіаною* (ЛГ). Таким чином, кожна багатовимірна гаусіана може бути представлена ЛГ. Зворотне твердження також є справедливим. Кожна МБ з лінійними УЙР являє собою нормальний сумісний розподіл.

У ЛГ формування висновку може виконуватись багатьма способами, один з них — просто перетворити ЛГ у багатовимірну гаусіану, над якою виконувати всі операції. Інший спосіб — адаптувати алгоритм виключення змінної або алгоритм об'єднаного дерева. Основна відмінність між підходами полягає у тому, що фактори не можуть бути більше представлені як таблиці, і не можуть бути представлені як гаусіани, тому що лінійні УЙР є не гаусіанами, а скоріше умовним розподілом, незважаючи на те, що кількість батьків дорівнює нулю.

Умовна лінійна гаусіана (УЛГ) — це МБ, яка містить як неперервні змінні Γ, так і дискретні змінні Δ, з такими обмеженнями:

— дискретний вузол не може мати неперервних батьків, таким чином, усі УЙР для дискретних вузлів можуть бути подані як у дискретних мережах Байєса;

— УЙР будь-якої неперервної змінної є лінійним УЙР, поданий будь-якою комбінацією дискретних батьків. Формально, якщо вузол Y має батьків $\{X_1,\dots,X_k\}\subseteq\Gamma$ і $D=\{D_1,\dots,D_l\}\subseteq\Delta$, він визначається як УЙР, з використанням таких параметрів: для кожного $d\in Dom(D)$, $\beta_{d,0},\dots,\beta_{d,k}$ і σ_d^2. УЙР тоді визначається як

$$P(Y\mid x,d)=N\left(Y;\ \beta_{d,0}+\sum_{i=1}^{k}\beta_{d,i}x_i,\sigma_d^2\right).$$

Це найбільш популярний вид гібридних моделей. Вони дозволяють лише лінійні відношення між неперервними змінними і не дозволяють дискретним змінним мати неперервних батьків. Широке використання таких моделей обумовлене їх математичною зручністю. В УЛГ задані значення дискретних змінних, розподіл неперервних змінних — багатовимірна гаусіана, отже сумісний ймовірнісний розподіл — це композиція гаусіан, з якими можна працювати, використовуючи аналітичні інструменти.

Якщо всі дискретні змінні задані, то умовні ймовірнісні розподіли неперервних змінних є лінійними умовними неперервними розподілами. Тоді при наданні будь-яких значень дискретним змінним умовна лінійна гаусіана редуцується у ЛГ і тому представляє нормальний розподіл. Звідси випливає, що сумісний розподіл, представлений УЛГ, є композицією гаусіан, де кожна композиція компонентів відповідає реалізації дискретних змінних.

Формування висновку у гібридних мережах Байєса.

Майже кожна система, яка використовує мережі Байєса, повинна виконувати ймовірнісний висновок, тобто знаходити розподіл певних змінних при деяких конкретних умовах. Для простих МБ розроблені точні і наближені методи формування висновку. Точні методи зазвичай дають точний результат, тоді як наближені методи намагаються отримати якомога ближчі вихідні дані. Точні методи застосовують для мереж, що відносяться до класу поодинці з'єднаних мереж, також відомих як полідерева. Мережа відноситься до цього класу, якщо основний ненаправлений граф має не більше одного шляху між двома вузлами. Коли мережа має багато з'єднань між вершинами, можна використати метод кластеризації для перетворення її у полідерево, а потім виконати точний висновок. На практиці існують мережі великого розміру, для яких формування точного висновку є неможливим, тому застосовують наближені методи формування висновку: відбору проб, вибору за значущістю, зваженої правдоподібності та інш. [3]. Для наближеного висновку найбільш відомим є стохастичний відбір проб, він застосовується для будь-якого класу гібридних моделей, не тільки УЛГ. Однак алгоритми за методом спроб можуть витрачати багато часу для наближення до достовірної відповіді і тому не підходять для задач у реальному часі.

Існують точні методи формування висновку для гібридних МБ, такі як алгоритм виключення змінної та клік-дерева.

Оскільки при обчисленні ймовірності змінних мережі задача стає NP-складною — з нелінійною поліноміальною оцінкою кількості ітерацій, для уникнення цієї проблеми використовують *алгоритм виключення змінної*. Ідея запобігання створенню експоненційно великих факторів полягає у виключенні деякої змінної, яка не з'являється у подальших обчисленнях.

Спершу потрібно перетворити умовні розподіли ймовірностей у набір факторів F. Оскільки F включає змінні, які не з'являються у послідовності, вибирається така змінна Y і вона виключається. Виключення виконується шляхом вибору усіх факторів, до яких належить Y, перемноженням їх, виключенням шляхом сумування Y і поверненням результату у F. Коли у F залишаться лише змінні запиту, усі згадані фактори перемножуються разом і результуючий фактор повертається як результат запиту.

Алгоритм виключення змінної

$P = \{P_1, ..., P_n\}$ — набір умовних ймовірнісних розподілів, Q — змінні запиту, f_i — фактор представлення P_i.

1. $F = \{f_1, ..., f_n\}$ — набір факторів, X — набір випадкових змінних, що використовуються у F.
2. $Y = X - Q$.
3. Поки $(Y \neq \varnothing)$.
4. Вибираємо деяку змінну $Y' \in Y$ і набір $Y = Y - \{Y'\}$.
5. Вилучаємо з F усі фактори $f_1 ... f_k$ з Y'.
6. $f = \prod_{i=1}^{k} f_i$, f^- буде результатом виключення y з f шляхом сумування.
7. Додамо f^- до F.
8. Кінець циклу поки.
9. Повертаємо $\prod_{f \in F} f$.

Складність алгоритма виключення змінних визначається розміром факторів, що виникають. Це залежить від обраного порядку виключення. На жаль, знаходження оптимального порядку виключення є NP-складною задачею навіть з використанням жадібних евристик для вибору порядку виключення. Більш того, у деяких мережах навіть оптимальний порядок виключення призводить до факторів, які є експонентою розміру МБ. У такому разі неможливо використати точні методи виведення і потрібно вдатися до наближення.

Найбільш популярним точним алгоритмом формування висновку для багатократно з'єднаних дискретних мереж Байєса є *алгоритм клік-дерева* (об'єднаного дерева або алгоритм кластеризації) [5]. Алгоритм клік-дерева перетворює початкову багатократно з'єднану мережу у ненаправлений поодинці з'єднаний граф, який називається клік-деревом. Спочатку в основному ненаправленому графі замінюють направлені ребра у мережі на ненаправлені ребра. Далі батьків спільного вузла з'єднують попарно, проводячи моралізацію (або "одруження") графа. Потім петлі у ненаправленому графі, які мають чотири або більше вузлів, розбивають на петлі, що містять щонайбільше три вузли. Іншими словами, вставляють деякі хорди у петлі і роблять таким чином ненаправлений граф хордальним. Цей процес називається тріангуляцією графа. Ненаправлений граф тріангулюється для породження малих кліків. Кліки формуються як фактори зв'язаних змінних з моралізованого, тріангульованого ненаправленого графа. Фактори над змінними, які містяться у кліку, називаються потенціалами. Кожне ребро асоціюється з фактором, який називається окремою множиною, над перетином змінних двох кліків, які він поєднує. Після побудови з мережі Байєса відповідного клік-дерева виконується алгоритм розповсюдження повідомлень між кліками (або алгоритм калібрації). Кожен клік відправляє повідомлення своєму сусідові після отримання повідомлення від усіх його інших сусідів, усього $2n-2$ повідомлень, якщо є n вузлів, з'єднаних у $n-1$ окремі множини. Слід зауважити, що у клік-дереві завжди можна відправити усі $2n-2$ повідомлення — наприклад, можна вибрати вузол як кореневий, відправити усі $n-1$ повідомлень до кореневого вузла і потім відправити $n-1$

повідомлення від кореня до листків. Відправка повідомлень до кореня називається *спрямованим вгору шляхом* або *шляхом проти течії*. Відправка повідомлень до листів називається *низхідним шляхом* або *шляхом вниз за течією*. Після виконання алгоритму калібрації кожний потенціал у дереві містить відповідний маргінальний розподіл над його змінними.

Спочатку необхідно розглянути, як обчислюється маргінальний розподіл певного кліка, що називається кореневим вузлом, а потім розширити алгоритм передачі повідомлення для обчислення ймовірнісного розподілу для усіх кліків.

Крок ініціалізації — поєднання умовних ймовірнісних розподілів у клік-дерево. Спочатку умовні ймовірнісні розподіли перетворюються у фактори і для кожного такого фактора вибирається клік C, який містить усі змінні у факторі. Потім ініціалізується потенціал кожного кліка як добуток умовних ймовірнісних розподілів, які асоціюються з ним. Якщо немає умовного ймовірнісного розподілу, який належить деякому кліку, він задається як потенціал, що дорівнює одиниці для кожного входу. Тепер уся потрібна інформація знаходиться у дереві і можна її використати для обчислення розподілу кореневого вузла.

Обчислення, пов'язані з алгоритмом виключення змінних, можуть бути подані у вигляді алгоритма передачі повідомлення, де повідомлення передаються по дереву, починаючи з листків і ідучи до кореня. Тобто, кожний вузол має відправити повідомлення своєму сусідові на шляху до кореня. А для цього він спочатку повинен отримати всі повідомлення від усіх інших сусідів, перемножити їх разом з їхніми потенціалами і виключити усі змінні, які не знаходяться у тому вузлі, якому відправляють повідомлення.

Алгоритм відправки повідомлення у клік-дереві від C_1 до C_2 вздовж множини S

$f_1 = \sum_{C_1 - S} \Phi(C_1)$ (Виключення шляхом сумування C_1).

$f_2 = \dfrac{f_1}{\Phi(S)}$ (Поділити на останнє відправлене повідомлення).

$\Phi(S) = f_1$ (Зберегти повідомлення).

$\Phi(C_2) = \Phi(C_2) \cdot f_2$ (Перемножити C_2 на повідомлення).

Алгоритм об'єднаного дерева (клік-дерева) можна адаптувати для гібридної мережі Байєса. З алгоритмічної точки зору, структурою алгоритма точного висновку в умовних лінійних гаусіанах є алгоритм [5], що базується на алгоритмі клік-дерева для дискретних мереж Байєса. В процесі адаптації алгоритма для умовних лінійних гаусіан постає питання: як представити потенціали кліків і окремі множини, і, зокрема, як представити функції над неперервними змінними. Обравши представлення потенціалів і окремих множин, необхідно підтвердити, що можна представити різні операції, необхідні для виконання алгоритма. За винятком маржиналізації, усі ці операції є природним узагальненням операцій, визначених для дискретних факторів, і операцій, визначених для канонічних форм. Маржиналізувати неперервні змінні можна, маржиналізуючи кожну канонічну форму.

Слід зауважити, що *передача повідомлень* не завжди можлива. Один з шляхів забезпечення упевненості, що передача повідомлень можлива — це створення об'єднаних дерев, які є сильнокореневими. У такому дереві можна вибрати клік R (що називається сильним коренем) такий, що не буде слабкої маржиналізації при відправці повідомлення назустріч R. Клік R у об'єднаному дереві T є *сильним коренем*, якщо кожні два суміжні кліки C_1 і C_2 такі, що C_1 ближчий до R, ніж C_2, задовольняють умові $(C_1 \cap C_2) \subseteq \Delta \vee (C_2 - C_1) \subseteq \Gamma$. Відсутність потреби виконувати слабку маржиналізацію на шляху вгору означає, що, коли відправляється повідомлення до кореня, там або немає неперервних змінних на окремій множині, або немає необхідності виключення будь−яких дискретних змінних шляхом сумування.

Коли дерево є сильнокореневим, можливе його калібрування з використанням алгоритма передачі повідомлень. Спочатку запускається шлях догори до R. Цей шлях є можливим, оскільки не включена слабка маржиналізація. Потім запускається шлях донизу. Цього разу виконується слабка маржиналізація, але вона буде завжди можливою, оскільки шляхом донизу завжди відправляється повідомлення від кліків, що вже отримали усі свої повідомлення, і тому представляють функцію ймовірності. Таким чином, канонічні форми мають представляти гаусіани і можуть бути згорнуті. Коли використовуються умовні фактори, $T = \varnothing$ для сильного кореня R і для будь-якого іншого кліка C. T містить неперервні змінні, які з'являються в окремій множині між C і його сусідами на шляху до R. Усі інші неперервні змінні у кліці є у H.

Дискретне свідчення може бути введене у дерево шляхом перемноження одного з потенціалів з відповідною індикаторною функцією, як у випадку дискретного об'єднаного дерева. Неперервне свідчення може бути введене шляхом реалізації кожного відповідного потенціалу, як у об'єднаному дереві для лінійних гаусіан. Свідчення може бути також введене після калібрування дерева, тоді неперервне свідчення теж має бути реалізовано у множинах.

Алгоритм калібрування сильно кореневого дерева

1. Вибір сильного кореня R.
2. Ініціалізація канонічних факторів у потенціалах і окремих множинах.
3. Перемноження потенціалів за допомогою умовних ймовірнісних розподілів.
4. Введення свідчення:

— дискретного: у будь-який потенціал;

— неперервного: в усі потенціали і окремі множини; видалення змінних, що спостерігаються, з дерева.

5. Відправка повідомлень догори до кореня R, використовуючи алгоритм відправки повідомлень для об'єднаних дерев.
6. Відправка повідомлень донизу до листів, використовуючи алгоритм відправки повідомлень для об'єднаних дерев.

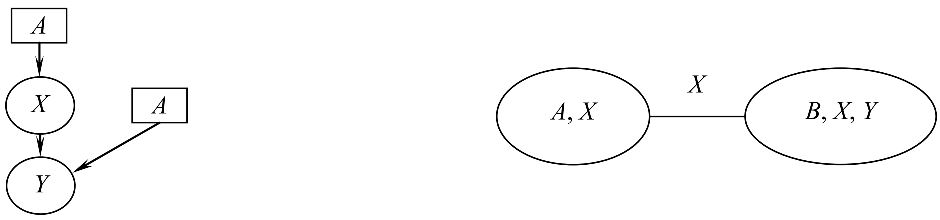
Однак, можна знайти об'єднане дерево, яке не є сильно кореневим, але у якому передача повідомлень все ще можлива. Наявність сильного кореня є достатньою, але не необхідною умовою для того, щоб операції алгоритма передачі повідомлень були добре визначеними у гібридному об'єднаному дереві.

Використання сильнокореневих дерев гарантує, що передача повідомлення не просто добре визначена, але також приводить до точних результатів. Якщо дерево не є сильнокореневим, алгоритм передачі повідомлень не приводить до правильних перших двох моментів і дерево не буде каліброваним, навіть якщо операції добре визначені. Фактично, передача повідомлень у деревах, які не є сильнокореневими, може призвести до невірних коваріаційних матриць.

Передача повідомлень у клік-деревах без сильного кореня призводить до невірної коваріайціної матриці, і зокрема до негативної дисперсії. Наприклад, у лінійній гаусіані умовні ймовірнісні розподіли дискретних змінних рівномірні і УЙРи неперервних змінних визначені таким чином (див. рисунок):

$$P(X \mid A) = \begin{cases} N(0,2) & A = 0, \\ N(0,6) & A = 1; \end{cases} \qquad P(Y \mid B, X) = \begin{cases} N(X, 0{,}01) & B = 0, \\ N(-X, 0{,}01) & B = 1. \end{cases}$$

Нехай використовується клік-дерево, яке не є сильно кореневим, і припустимо, що є свідчення $Y = 4$ (див. рисунок).



а *б*

Проста умовна лінійна гаусіана (а) та клік-дерево без сильного кореня (б)

Нехай алгоритм передачі повідомлень спершу відправляє повідомлення від кліка $\{A, X\}$ до кліка $\{B, X, Y\}$. Оскільки кліки поділяють одну змінну X, згортається апріорний розподіл

X і отримується повідомлення $M_1 = N(X;\ 0; 4)$. Коли перемножуються потенціали у $\{B, X, Y\}$ за допомогою цього повідомлення, отримується композиція двох гаусіан з однаковими вагами:

$$N(X,Y;\begin{pmatrix}0\\0\end{pmatrix},\begin{bmatrix}4 & 4\\4 & 4{,}01\end{bmatrix}, \text{коли } B=0;\ N(X,Y;\begin{pmatrix}0\\0\end{pmatrix},\begin{bmatrix}4 & -4\\-4 & 4{,}01\end{bmatrix}, \text{коли } B=1.$$

Після реалізації свідчення $Y = 4$ отримується нова композиція з двох гаусіан:

$$N(X;\ 3{,}99;\ 0{,}00998), \text{ коли } B = 0 \text{ і } N(X;\ -3{,}99;\ 0{,}00998), \text{ коли } B = 1.$$

Вага гаусіан у новій композиції буде тією самою, тому що Y має той самий маргінал, що у початковій суміші $N(Y;\ 0;\ 4{,}01)$ для обох $B = 0$ і $B = 1$, і відповідно свідчення має одну й ту ж правдоподібність для двох гаусіан.

Тепер необхідно відправити повідомлення назад до кліка $\{A, X\}$. Для цього необхідно зробити згортку двох гаусіан. Результатом згортки є повідомлення $M_2 = N(X;\ 0;\ 15{,}93)$. Слід зауважити, що у цьому прикладі свідчення спричинило збільшення дисперсії X.

Для включення повідомлення M_2 у клік $\{A, X\}$ необхідно поділити його за допомогою повідомлення M_1. Потім кожний вхід у потенціалі $\{A, X\}$ перемножується на результат ділення, тобто на $\frac{M_2}{M_1}$. Зокрема, для $A = 1$ необхідно виконати операцію: $\frac{N(0;\ 6)\cdot N(0;\ 15{,}93)}{N(0;\ 4)}$. Ця операція може бути виконана у канонічних формах $C(K, h, g)$, і для даного прикладу достатньо обчислити лише коефіцієнт K другого порядку. Необхідно нагадати, що взагалі $K = \Sigma^{-1}$ і тому для $\sigma^2 = 6$ $K = \frac{1}{6}$, для $\sigma^2 = 15{,}93$ $K = \frac{1}{15{,}93}$ і для $\sigma^2 = 4$ $K = \frac{1}{4}$. Таким чином, K другого порядку дорівнює: $K = \frac{1}{6} + \frac{1}{15{,}93} - \frac{1}{4} = -0{,}0206$. Однак, $K < 0$ не представляє легальну гаусіану — якщо використовується інверсна трансформація, $\sigma^2 = \frac{1}{-0{,}0206}$, що не є легальною дисперсію. Таким чином, виконувати передачу повідомлення у клік-дереві, що не є сильно кореневим, небезпечно.

Розглянуті особливості гібридних мереж Байєса показали, що гібридні мережі є узагальненням звичайних МБ для випадків, коли змінні задачі мають неперервний розподіл або дуже велику кількість можливих значень. Незважаючи на складність представлення та формування висновку у мережі, гібридні мережі доцільно використовувати для задач з великими обсягами статистичної інформації, наприклад для фінансових задач. Для гібридних мереж можна застосовувати ті самі точні алгоритми формування висновку (алгоритм виключення змінної, передачі повідомлення, клік-дерева і т.п.), адаптуючи їх. Точний висновок в умовних лінійних гаусіанах складніший, ніж у дискретних мережах Байєса. Тому не дивно, що інколи приходиться вдаватися до наближеного алгоритма формування висновку. При цьому необхідно балансувати між необхідністю точного висновку і обчислювальною складністю задачі, оскільки навіть для простих структур мережі, коли точний висновок в дискретних мережах може бути сформований за лінійний час, навіть наближений висновок у гібридних моделях є NP-складною задачею. Отже, інколи є сенс використовувати дискретизацію неперервних змінних при формуванні висновку у гібридних мережах.

Література

1. Бідюк, П.І. Основні етапи побудови і приклади застосування мереж Байєса / П.І. Бідюк, Н.В. Кузнєцова // Систем. дослідж. та інформ. технології. — 2007. — № 4. — С. 26 — 39.
2. Кузнєцова, Н.В. Системний підхід до аналізу кредитних ризиків з використанням мереж Байєса / Н.В. Кузнєцова, П.І. Бідюк // Наук. вісті НТУУ "КПІ". — 2008. — № 3. — С. 11 — 24.

3. Gogate, V. Approximate Inference Algorithms for Hybrid Bayesian Networks with Discrete Constraints / V. Gogate, R. Dechter // UAI. — 2005. — P. 209 — 216.
4. Pearl, J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference / J. Pearl. — San Mateo, CA (USA): Morgan Kauffmann Publishers, Inc., 1988. — 550 p.
5. Lauritzen, S.L. Local computations with probabilities on graphical structures and their application to expert systems / S.L. Lauritzen, D.J. Spiegelhalter // J. of the Royal Statist. Soc., ser. B. — 1988. — Vol. 50, № 2. — P. 157 — 224.

Рецензент д-р техн. наук, проф. Одес. нац. політехн, ун-ту Петрушин В.С.

Надійшла до редакції 30 березня 2009 р.