鈴木譲「ベイジアンネットワーク入門」(培風館)
4. 確率的推論
4.1 確率分布の計算

鈴木譲

大阪大学

2009 年 1 月 14 日 (木)

あらまし

確率的推論

$X_i(\Omega) < \infty,\ i \in \mathcal{N} := \{1, \cdots, N\}$
$P_{X_1 \cdots X_N}(x_1, \cdots, x_N)$: $x_1 \in X_1(\Omega), \cdots, x_N \in X_N(\Omega)$ の確率

確率的推論

$$P_{X_i}(x_i) := \sum_{x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N} P_{X_1 \cdots X_N}(x_1, \cdots, x_N) \tag{1}$$

を計算 (N とともに指数的な計算が必要)

確率変数 2 個以上でも同様:

$$P_{X_i X_j}(x_i, x_j) := \sum_{x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{j-1}, x_{j+1}, \cdots, x_N} P_{X_1 \cdots X_N}(x_1, \cdots, x_N)$$

例 4.1: 銃による殺人事件で 3 人が容疑

$X_1(\Omega) = X_2(\Omega) = \{1, 2, 3\}$
X_0: 裁判所での証言
X_1: 犯人が誰であったか
X_2: 指紋が誰のものであったか
X_3 で指紋の調査結果

$$P_{X_2|X_1}(\cdot|\cdot) := \begin{bmatrix} P_{X_2|X_1}(1|1) & P_{X_2|X_1}(2|1) & P_{X_2|X_1}(3|1) \\ P_{X_2|X_1}(1|2) & P_{X_2|X_1}(2|2) & P_{X_2|X_1}(3|2) \\ P_{X_2|X_1}(1|3) & P_{X_2|X_1}(2|3) & P_{X_2|X_1}(3|3) \end{bmatrix}$$
$$= \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}$$

$$X_0 \longrightarrow X_1 \longrightarrow X_2 \longrightarrow X_3$$

例 4.1: 証言 $X_0 = e_0$ が得られると

$$P_{X_1|X_0}(\cdot|e_0) := [P_{X_1|X_0}(1|e_0), P_{X_1|X_0}(2|e_0), P_{X_1|X_0}(3|e_0)] = [p_1, p_2, p_3]$$

$$\begin{aligned}
P_{X_2|X_0}(\cdot|e_0) &\propto P_{X_1|X_0}(\cdot|e_0) \times P_{X_2|X_1}(\cdot|\cdot) \\
&= [p_1, p_2, p_3] \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \\
&= \left[\sum_{i=1}^{3} p_i q_{i1}, \sum_{i=1}^{3} p_i q_{i2}, \sum_{i=1}^{3} p_i q_{i3} \right]
\end{aligned}$$

$$X_0 \longrightarrow X_1 \longrightarrow X_2 \longrightarrow X_3$$

$X_0 = e_0$

例 4.1: 指紋調査の結果が届く前

$$P_{X_3|X_2}(e_3|\cdot) := [P_{X_3|X_2}(e_3|1), P_{X_3|X_2}(e_3|2), P_{X_3|X_2}(e_3|3)] \propto [1, 1, 1]$$

$$\begin{aligned} P_{X_2|X_0,X_3}(\cdot|e_0, e_3) &\propto P_{X_2|X_0}(\cdot|e_0) \times P_{X_3|X_2}(e_3|\cdot) \\ &= \left[\sum_{i=1}^{3} p_i q_{i1} \times 1, \sum_{i=1}^{3} p_i q_{i2} \times 1, \sum_{i=1}^{3} p_i q_{i3} \times 1\right] \end{aligned}$$

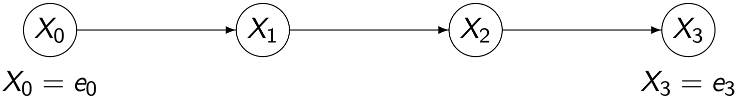
$$\begin{tikzcd} X_0 \arrow[r] & X_1 \arrow[r] & X_2 \arrow[r] & X_3 \end{tikzcd}$$

$X_0 = e_0$ $\qquad\qquad$ $X_3 = e_3$

例 4.1: 指紋調査の結果が届くと

$$P_{X_3|X_2}(e_3|\cdot) \propto [r_1, r_2, r_3]$$

$$\begin{aligned} P_{X_2|X_0X_3}(\cdot|e_0, e_3) &\propto P_{X_2|X_0}(\cdot|e_0) \times P_{X_3|X_2}(e_3|\cdot) \\ &= \left[\sum_{i=1}^{3} p_i q_{i1} \times r_1, \sum_{i=1}^{3} p_i q_{i2} \times r_2, \sum_{i=1}^{3} p_i q_{i3} \times r_3\right] \end{aligned}$$



例 4.1: 指紋調査の結果が届くと (続)

$$P_{X_3|X_1}(e_3|\cdot) \propto \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{3} q_{1j} r_j \\ \sum_{j=1}^{3} q_{2j} r_j \\ \sum_{j=1}^{3} q_{3j} r_j \end{bmatrix}$$

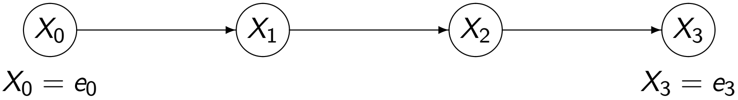
$$\begin{aligned} P_{X_1|X_0X_3}(\cdot|e_0, e_3) &\propto P_{X_3|X_1}(e_3|\cdot) \times P_{X_1|X_0}(\cdot|e_0) \\ &= \left[\sum_{j=1}^{3} q_{1j} r_j \times p_1, \sum_{j=1}^{3} q_{2j} r_j \times p_2, \sum_{j=1}^{3} q_{3j} r_j \times p_3 \right], \end{aligned}$$

例 4.1: 容疑者 1 に強いアリバイがあることわかると

$$P_{X_1|X_0}(\cdot|e_0) = [p_1, p_2, p_3] \to P_{X_1|X_0}(\cdot|e_0) = [p'_1, p'_2, p'_3]$$

$$P_{X_2|X_0}(\cdot|e_0) = [p'_1, p'_2, p'_3] \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}$$

$$= \left[\sum_{i=1}^{3} p'_i q_{i1}, \sum_{i=1}^{3} p'_i q_{i2}, \sum_{i=1}^{3} p'_i q_{i3} \right]$$



例 4.1: 容疑者 1 に強いアリバイがあることわかると (続)

$$P_{X_1|X_0X_3}(\cdot|e_0,e_3) \propto P_{X_1|X_0}(\cdot|e_0) \times P_{X_3|X_1}(e_3|\cdot)$$
$$= \left[p'_1 \times \sum_{j=1}^{3} q_{1j}r_j, p'_2 \times \sum_{j=1}^{3} q_{2j}r_j, p'_3 \times \sum_{j=1}^{3} q_{3j}r_j\right],$$
$$P_{X_2|X_0X_3}(\cdot|e_0,e_3) \propto P_{X_2|X_0}(\cdot|e_0) \times P_{X_3|X_2}(e_3|\cdot)$$
$$= \left[\sum_{i=1}^{3} p'_i q_{i1} \times r_1, \sum_{i=1}^{3} p'_i q_{i2} \times r_2, \sum_{i=1}^{3} p'_i q_{i3} \times r_3\right]$$

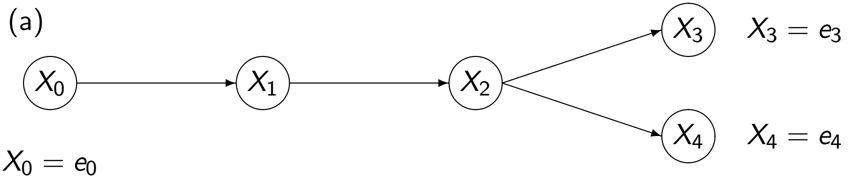
例 4.2: (1) 別の指紋研究所での調査結果 $X_4 = e_4$

$$P_{X_4|X_2}(e_4|\cdot) := [P_{X_4|X_2}(e_4|1), P_{X_4|X_2}(e_4|2), P_{X_4|X_2}(e_4|3)] \propto [s_1, s_2, s_3]$$

$$P_{X_3X_4|X_2}(e_3, e_4|\cdot) := [P(e_3, e_4|1), P(e_3, e_4|2), P(e_3, e_4|3)] \propto [r_1 s_1, r_2 s_2, r_3 s_3]$$

$$P_{X_2|X_0X_3X_4}(\cdot|e_0, e_3, e_4)$$
$$\propto \left[\sum_{i=1}^{3} p'_i q_{i1} \times r_1 s_1, \sum_{i=1}^{3} p'_i q_{i2} \times r_2 s_2, \sum_{i=1}^{3} p'_i q_{i3} \times r_3 s_3\right]$$

(a)



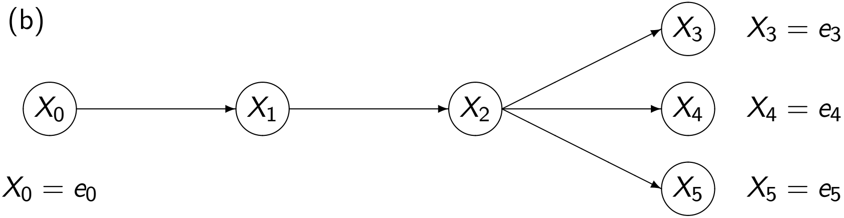
例 4.2: (2) 容疑者 2 が「指紋は私のものです」と主張して、$X_5 = e_5$

$$P_{X_5|X_2}(e_5|\cdot) := [P_{X_5|X_2}(e_5|1), P_{X_5|X_2}(e_5|2), P_{X_5|X_2}(e_5|3)] \propto [0,1,0]$$

$$P_{X_3X_4X_5|X_2}(e_3, e_4, e_5|\cdot) \quad \propto \quad [r_1 s_1 \times 0, r_2 s_2 \times 1, r_3 s_3 \times 0]$$

$$P_{X_3X_4X_5|X_2}(e_3, e_4, e_5|\cdot) = [0,1,0]$$

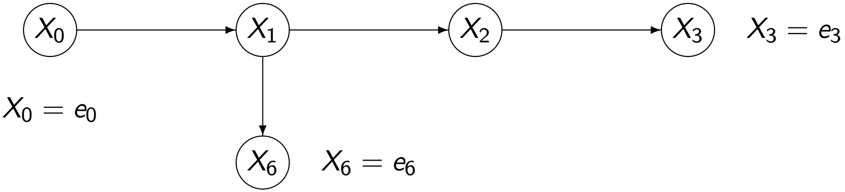
(b)



例 4.2: (3) 容疑者 1 のアリバイ $X_6 = e_6$ という証拠

$$P_{X_6|X_1}(e_6|\cdot) := [P_{X_6|X_1}(e_6|1), P_{X_6|X_1}(e_6|2), P_{X_6|X_1}(e_6|3)] \propto [t_1, t_2, t_3]$$

(c)



例 4.2: (3) 容疑者 1 のアリバイ $X_6 = e_6$ という証拠 (続)

$$
\begin{aligned}
P_{X_3X_6|X_1}(e_3, e_6|\cdot) &\propto P_{X_3|X_1}(e_3|\cdot) \times P_{X_6|X_1}(e_6|\cdot) \\
&= \left[\sum_{j=1}^{3} q_{1j} r_j \times t_1, \sum_{j=1}^{3} q_{2j} r_j \times t_2, \sum_{j=1}^{3} q_{3j} r_j \times t_3\right]
\end{aligned}
$$

$$
\begin{aligned}
& P_{X_1|X_0X_3X_6}(\cdot|e_0, e_3, e_6) \\
\propto\ & P_{X_1|X_0}(\cdot|e_0) \times P_{X_3X_6|X_1}(e_3, e_6|\cdot) \\
=\ & \left[\sum_{j=1}^{3} t_1 q_{1j} r_j \times p'_1, \sum_{j=1}^{3} t_2 q_{2j} r_j \times p'_2, \sum_{j=1}^{3} t_3 q_{3j} r_j \times p'_3\right]
\end{aligned}
$$

因子グラフ

$\mathcal{M}$: 有限集合
$\mathcal{N}(a) \subseteq \mathcal{N},\ a \in \mathcal{M}$

$$P_{X_1 \cdots X_N}(x_1, \cdots, x_N) = \frac{1}{Z} \prod_{a \in \mathcal{M}} f_a(x_i, i \in \mathcal{N}(a)), \tag{2}$$

$$Z := \sum_{x_i, i \in \mathcal{N}} \prod_{a \in \mathcal{M}} f_a(x_i, i \in \mathcal{N}(a))$$

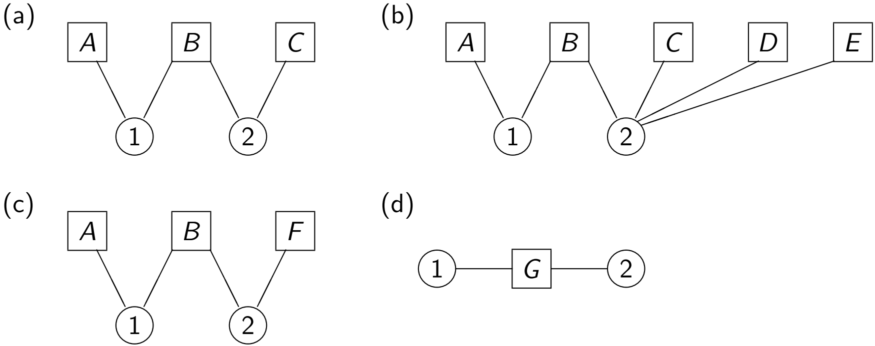
- 変数頂点 $\mathcal{N} = \{1, \cdots, N\}$ の要素
- 因子頂点 $\mathcal{M}$ の要素
- 辺 $\{(i, a) \mid i \in \mathcal{N}(a), a \in \mathcal{M}\}$ の要素

$i \in \mathcal{N}$ と辺で結ばれた因子頂点の集合を $\mathcal{M}(i) \subseteq \mathcal{M}$ であらわすと、

$$\{(i, a) \mid i \in \mathcal{N}(a), a \in \mathcal{M}\} = \{(i, a) \mid i \in \mathcal{N}, a \in \mathcal{M}(i)\}$$

因子グラフ (続)

(2) を、変数頂点○、因子頂点□、およびそれらを結ぶ辺で表す



因子グラフの例

(a)

$$
\begin{aligned}
& P_{X_0X_1X_2X_3}(e_0, x_1, x_2, e_3) \\
= \; & P_{X_0}(e_0)P_{X_1|X_0}(x_1|e_0)P_{X_2|X_3}(x_2|x_3)P_{X_3|X_2}(e_3|x_2) \\
= \; & f_A(x_1)f_B(x_1, x_2)f_C(x_2)
\end{aligned}
$$

$\mathcal{N} = \{1, 2\}$, $\mathcal{M} = \{A, B, C\}$, $\mathcal{N}(A) = \{1\}$, $\mathcal{N}(B) = \{1, 2\}$,
$\mathcal{N}(C) = \{1, 2\}$

(b)

$$
\begin{aligned}
& P_{X_0X_1X_2X_3X_4X_5}(e_0, x_1, x_2, x_3, x_4, x_5) \\
= \; & P_{X_0}(e_0)P_{X_1|X_0}(x_1|e_0)P_{X_2|X_3}(x_2|x_3)P_{X_3|X_2}(e_3|x_2) \\
& \cdot P_{X_4|X_2}(e_4|x_2)P_{X_5|X_2}(e_5|x_2) \\
= \; & f_A(x_1)f_B(x_1, x_2)f_C(x_2)f_D(x_2)f_E(x_2)
\end{aligned}
$$

$\mathcal{N} = \{1, 2\}, \mathcal{M} = \{A, B, C, D, E\}$,
$\mathcal{N}(A) = \{1\}, \mathcal{N}(B) = \{1, 2\}, \mathcal{N}(C) = \mathcal{N}(D) = \mathcal{N}(E) = \{2\}$

因子グラフの例 (続)

(4) で $f_F(x_2) := f_C(x_2)f_D(x_2)f_E(x_2)$ とおくと、

$$P_{X_0X_1X_2X_3X_4X_5}(e_0, x_1, x_2, x_3, x_4, x_5) = f_A(x_1)f_B(x_1, x_2)f_F(x_2)$$

$f_G(x_1, x_2) := f_A(x_1)f_B(x_1, x_2)f_F(x_2)$ とおくと、

$$P_{X_0X_1X_2X_3X_4X_5}(e_0, x_1, x_2, x_3, x_4, x_5) = f_G(x_1, x_2)$$

巡回経路

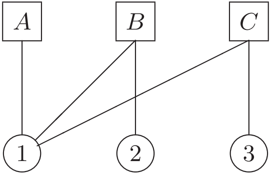
巡回経路

異なる 2 因子頂点を結ぶ辺の列が 2 個以上存在

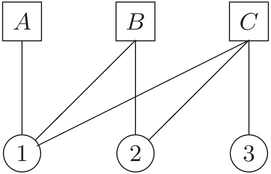
$$P_{X_1}(x_1)P_{X_2|X_1}(x_2|x_1)P_{X_3|X_2}(x_3|x_2) = f_A(x_1)f_B(x_1,x_2)f_C(x_2,x_3)$$

$$P_{X_1}(x_1)P_{X_2|X_1}(x_2|x_1)P_{X_3|X_1,X_2}(x_3|x_1,x_2) = f_A(x_1)f_B(x_1,x_2)f_C(x_1,x_2,x_3)$$

(a)



(b)



アルゴリズム 4.1

$i \in \mathcal{N}$, $a \in \mathcal{M}$
$n_{i\to a}$, $m_{a\to i}$: $\mathcal{X}_i$ から非負実数への写像

$$n_{i\to a}(x_i) := 1,\ m_{a\to i}(x_i) := 1,\ x_i \in \mathcal{X}_i$$

$$n_{i\to a}(x_i) := \prod_{c\in\mathcal{M}(i)\setminus\{a\}} m_{c\to i}(x_i), \tag{5}$$

$$m_{a\to i}(x_i) := \sum_{\mathbf{x}_{a,i}\in\mathcal{X}_{a,i}} f_a(\mathbf{x}_a) \prod_{j\in\mathcal{N}(a)\setminus\{i\}} n_{j\to a}(x_j) \tag{6}$$

アルゴリズム 4.1

$n_{i\to a}(x_i) := 1$, $m_{a\to i}(x_i) := 1$, $x_i \in \mathcal{X}_i$ とおき、各 $(i,a) \in \mathcal{N}\times\mathcal{M}$ に対して，毎回同時に (5), (6) を適用して，$n_{i\to a}(x_i)$, $m_{a\to i}(x_i)$ を更新

アルゴリズム 4.1 (続)

(5) より、$\{m_{a\to i}(x_i)\}$ だけで (5),(6) を更新してもよい:

$$m_{a\to i}(x_i) := \sum_{\mathbf{x}_{a,i}\in\mathcal{X}_{a,i}} f_a(\mathbf{x}_a) \prod_{j\in\mathcal{N}(a)\backslash\{i\}} \prod_{c\in\mathcal{M}(j)\backslash\{a\}} m_{c\to j}(x_j) \tag{7}$$

メッセージ

$\{n_{i\to a}(x_i)\}$, $\{m_{a\to i}(x_i)\}$, $x_i \in \mathcal{X}_i$

定理 4.1

定理 4.1

因子グラフが巡回経路を含まないとき、アルゴリズム 1 において (5), (6) の有限回の適用で $\{n_{i\to a}(x_i)\}$, $\{m_{a\to i}(x_i)\}$ が一定値に収束

その収束値について、

$$b_i(x_i) \propto \prod_{a\in N(i)} m_{a\to i}(x_i), \tag{8}$$

$$\sum_{x_i\in\mathcal{X}_i} b_i(x_i) = 1 \tag{9}$$

を満足する $b_i(x_i)$ は，$P_{X_i}(x_i)$ に等しい。

定理 4.1 (続)

$$b_a(\mathbf{x}_a) \propto f_a(\mathbf{x}_a) \prod_{i \in N(a)} n_{i \to a}(x_i), \tag{10}$$

$$\sum_{\mathbf{x}_a \in \mathcal{X}_a} b_a(\mathbf{x}_a) = 1 \tag{11}$$

を満足する $b_a(\mathbf{x}_a)$ は，$P_{\mathbf{X}_a}(\mathbf{x}_a) := P_{X_j, j \in \mathcal{N}(a)}(x_j, j \in \mathcal{N}(a))$ に等しい。

$$\sum_{\mathbf{x}_{a,i} \in \mathcal{X}_{a,i}} b_a(\mathbf{x}_a) = b_i(x_i) \tag{12}$$

ビリーフ

$\{b_i(x_i)\}_{i \in N}$, $\{b_a(\mathbf{x}_a)\}_{a \in M}$

ビリーフの値が更新されることを**ビリーフプロパゲーション**という

例 4.5

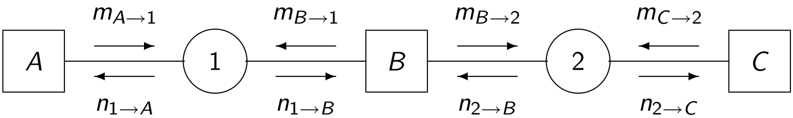
例 4.1 で

$f_A(x_1) := P_{X_0}(e_0)P_{X_1|X_0}(x_1|e_0)$

$f_B(x_1, x_2) := P_{X_2|X_1}(x_2|x_1)$

$f_C(x_2) := P_{X_3|X_2}(e_3|x_2)$

(a)



例 4.5 (続)

$$
\begin{aligned}
n_{2\to B}(x_2) &= m_{C\to 2}(x_2),\\
n_{1\to A}(x_1) &= m_{B\to 1}(x_1),\\
n_{1\to B}(x_1) &= m_{A\to 1}(x_1),\\
n_{2\to C}(x_2) &= m_{B\to 2}(x_2),
\end{aligned}
$$

$$
\begin{aligned}
m_{A\to 1}(x_1) &= f_A(x_1) = P_{X_0}(e_0)P_{X_1|X_0}(x_1|e_0),\\
m_{B\to 2}(x_2) &= \sum_{x_1} f_B(x_1,x_2)n_{1\to B}(x_1) = \sum_{x_1} f_B(x_1,x_2)m_{A\to 1}(x_1)\\
&= \sum_{x_1} f_A(x_1)f_B(x_1,x_2)\\
&= \sum_{x_1} P_{X_0}(e_0)P_{X_1|X_0}(x_1|e_0)P_{X_2|X_1}(x_2|x_1),
\end{aligned}
$$

例 4.5 (続)

$$
\begin{aligned}
m_{C\to 2}(x_2) &= f_C(x_2) = P_{X_3|X_2}(e_3|x_2), \\
m_{B\to 1}(x_1) &= \sum_{x_2} f_B(x_1, x_2) n_{2\to B}(x_2) = \sum_{x_2} f_B(x_1, x_2) m_{C\to 2}(x_1) \\
&= \sum_{x_2} f_B(x_1, x_2) f_C(x_2) = \sum_{x_2} P_{X_2|X_1}(x_2|x_1) P_{X_3|X_2}(e_3|x_2),
\end{aligned}
$$

$$
\begin{aligned}
b_1(x_1) &\propto m_{A\to 1}(x_1) m_{B\to 1}(x_1) \\
&= P_{X_0}(e_0) P_{X_1|X_0}(x_1|e_0) \sum_{x_2} P_{X_2|X_1}(x_2|x_1) P_{X_3|X_2}(e_3|x_2) \\
b_2(x_2) &\propto m_{B\to 2}(x_2) m_{C\to 2}(x_2) \\
&= \sum_{x_1} P_{X_0}(e_0) P_{X_1|X_0}(x_1|e_0) P_{X_2|X_1}(x_2|x_1) P_{X_3|X_2}(e_3|x_2)
\end{aligned}
$$

例 4.5 (続)

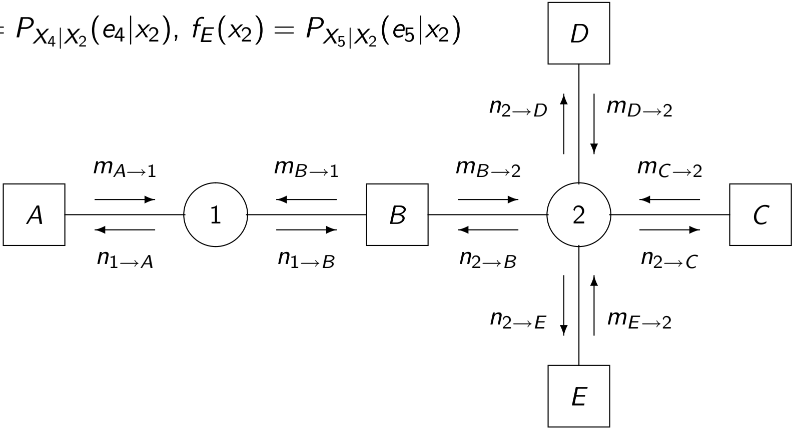
$$
\begin{aligned}
b_A(x_A) &\propto f_A(x_1) n_{1\to A}(x_1) \\
&= P_{X_0}(e_0) P_{X_1|X_0}(x_1|e_0) \sum_{x_2} P_{X_2|X_1}(x_2|x_1) P_{X_3|X_2}(e_3|x_2), \\
b_B(x_B) &\propto f_B(x_1, x_2) n_{1\to B}(x_1) n_{2\to B}(x_2) \\
&= P_{X_0}(e_0) P_{X_1|X_0}(x_1|e_0) P_{X_2|X_1}(x_2|x_1) P_{X_3|X_2}(e_3|x_2), \\
b_C(x_C) &\propto f_C(x_2) n_{2\to C}(x_2) \\
&= \sum_{x_1} P_{X_0}(e_0) P_{X_1|X_0}(x_1|e_0) P_{X_2|X_1}(x_2|x_1) P_{X_3|X_2}(e_3|x_2).
\end{aligned}
$$

例 4.5 (続)

例 4.2

$f_D(x_2) = P_{X_4|X_2}(e_4|x_2)$, $f_E(x_2) = P_{X_5|X_2}(e_5|x_2)$

(b)



例 4.5 (続)

$$m_{D\to 2}(x_2) = f_D(x_2) = P_{X_4|X_2}(e_4|x_2),$$
$$m_{E\to 2}(x_2) = f_E(x_2) = P_{X_5|X_2}(e_5|x_2),$$

$$\begin{aligned} n_{2\to B}(x_2) &= m_{C\to 2}(x_2)m_{D\to 2}(x_2)m_{E\to 2}(x_2)f_C(x_2)f_D(x_2)f_E(x_2) \\ &= P_{X_3|X_2}(e_3|x_2)P_{X_4|X_2}(e_4|x_2)P_{X_5|X_2}(e_5|x_2), \\ n_{2\to C}(x_2) &= n_{2\to D}(x_2) = n_{2\to E}(x_2) = m_{B\to 2}(x_2), \end{aligned}$$

$$\begin{aligned} b_2(x_2) &\propto m_{B\to 2}(x_2)m_{C\to 2}(x_2)m_{D\to 2}(x_2)m_{E\to 2}(x_2) \\ &= \sum_{x_1} P_{X_0}(e_0)P_{X_1|X_0}(x_1|e_0)P_{X_2|X_1}(x_2|x_1) \\ &\qquad \cdot P_{X_3|X_2}(e_3|x_2)P_{X_4|X_2}(e_4|x_2)P_{X_5|X_2}(e_5|x_2) \end{aligned}$$

アルゴリズム 4.2

定理 4.1 は，アルゴリズム 4.2 でも成立する:
$n_{i\to a}(x_i) := 1,\ m_{a\to i}(x_i) := 1,\ x_i \in \mathcal{X}_i$

アルゴリズム 4.2

$(i,a) \in \mathcal{N} \times \mathcal{M}$ に対して事前に更新する順序を決めておく。1 回に 1 つの (i,a) について (5), (6) を適用して，$n_{i\to a}(x_i), m_{a\to i}(x_i)$ を更新

巡回経路を含むとき、定理 4.2 は成立しない

- 収束しない
- 収束しても，$b_a = P_{X_a}, b_i = P_{X_i}$ などが成立しない

例: $0 < \epsilon < 0.125$ として，

$$P_{X_1X_2X_3}(x_1, x_2, x_3) = f_A(x_1, x_2) f_B(x_1, x_3) f_C(x_2, x_3),$$

$$f_A(x_1, x_2) = \begin{cases} 1-\epsilon, & x_1 = x_2 \\ \epsilon, & x_1 \neq x_2, \end{cases}$$

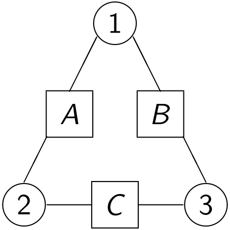
$$f_B(x_1, x_3) = \begin{cases} 1-\epsilon, & x_1 = x_3 \\ \epsilon, & x_1 \neq x_3, \end{cases}$$

$$f_C(x_2, x_3) = \begin{cases} \epsilon, & x_2 = x_3 \\ 1-\epsilon, & x_2 \neq x_3 \end{cases}$$

巡回経路を含むとき、定理 4.2 は成立しない (続)

$$\begin{aligned} b_A(x_1, x_2) &= P_{X_1X_2}(x_1, x_2) \\ b_B(x_1, x_3) &= P_{X_1X_3}(x_1, x_3) \\ b_C(x_2, x_3) &= P_{X_2X_3}(x_2, x_3) \end{aligned}$$

が同時に成立しない。



巡回経路を含むとき、定理 4.2 は成立しない (続)

$m_{A\to 1}(x_1) = m_{B\to 1}(x_1) = 1,\ m_{A\to 2}(x_2) = m_{C\to 2}(x_2) = 1,$
$m_{B\to 3}(x_3) = m_{C\to 3}(x_3) = 1,\ x_1 \in \mathcal{X}_1,\ x_2 \in \mathcal{X}_2,\ x_3 \in \mathcal{X}_3$

$$m_{A\to 1}(x_1) = \sum_{x_2\in\mathcal{X}_2} f_A(x_1, x_2) m_{C\to 2}(x_2) = 1$$

となり、すべての $a = A, B, C$, $i = 1, 2, 3$ で $m_{a\to i}(x_i)$ の値が変更されることなく，$b_i(x_i) = \frac{1}{2}$, $i = 1, 2, 3$ に収束

$$\begin{aligned} b_A(x_1, x_2) &= P_{X_1X_2}(x_1, x_2) \\ b_B(x_1, x_3) &= P_{X_1X_3}(x_1, x_3) \\ b_C(x_2, x_3) &= P_{X_2X_3}(x_2, x_3) \end{aligned}$$

が同時に成立しない。

巡回経路を含むとき、定理 4.2 は成立しない (続)

$$b_A(x_1,x_2) = f_A(x_1,x_2)m_{B\to 1}(x_1)m_{C\to 2} = f_A(x_1,x_2), \tag{14}$$
$$b_B(x_1,x_3) = f_B(x_1,x_3)m_{A\to 1}(x_1)m_{C\to 3} = f_B(x_1,x_3), \tag{15}$$
$$b_C(x_2,x_3) = f_C(x_2,x_3)m_{A\to 2}(x_2)m_{B\to 3} = f_C(x_2,x_3) \tag{16}$$

$$\begin{aligned} P_{X_1X_2X_3}(0,1,0) + P_{X_1X_2X_3}(0,1,1) &= \epsilon, \\ P_{X_1X_2X_3}(1,0,0) + P_{X_1X_2X_3}(1,0,1) &= \epsilon, \end{aligned}$$

$$\begin{aligned} P_{X_1X_2X_3}(0,0,1) + P_{X_1X_2X_3}(0,1,1) &= \epsilon, \\ P_{X_1X_2X_3}(1,0,0) + P_{X_1X_2X_3}(1,1,0) &= \epsilon, \end{aligned}$$

$$\begin{aligned} P_{X_1X_2X_3}(0,0,0) + P_{X_1X_2X_3}(1,0,0) &= \epsilon, \\ P_{X_1X_2X_3}(0,1,1) + P_{X_1X_2X_3}(1,1,1) &= \epsilon \end{aligned}$$

特に $P_{X_1X_2X_3}(x_1,x_2,x_3) \le \epsilon < 0.125$ のとき、

$$\sum_{x_1\in\mathcal{X}_1,x_2\in\mathcal{X}_2,x_3\in\mathcal{X}_3} P_{X_1X_2X_3}(x_1,x_2,x_3) < 1$$

最大事後確率設定

最大事後確率設定

$$P_{X_1\cdots X_N}(x_1,\cdots,x_N) := \max_{x_1,\cdots,x_N} P_{X_1\cdots X_N}(x_1,\cdots,x_N)$$

を最大にする $x_1 \in \mathcal{X}_1, \cdots, x_N \in \mathcal{X}_N$ を見いだす

$$n'_{i\to a}(x_i) := 1,\ \ m'_{a\to i}(x_i) := 1,\ \ x_i \in \mathcal{X}_i$$

$$n'_{i\to a}(x_i) := \prod_{c\in\mathcal{M}(i)\backslash\{a\}} m'_{c\to i}(x_i), \tag{17}$$

$$m'_{a\to i}(x_i) := \max_{\mathbf{x}_{a,i}\in\mathcal{X}_{a,i}} f_a(\mathbf{x}_a) \prod_{j\in\mathcal{N}(a)\backslash\{i\}} n'_{j\to a}(x_j) \tag{18}$$

アルゴリズム 4.3

$n'_{i\to a}(x_i) := 1$, $m'_{a\to i}(x_i) := 1$, $x_i \in \mathcal{X}_i$ とおき、$(i,a) \in \mathcal{N}\times\mathcal{M}$ に対して, 毎回同時に (17), (18) を適用して, $n'_{i\to a}(x_i)$, $m'_{a\to i}(x_i)$ を更新

定理 4.3

定理 4.3

因子グラフが巡回経路を含まないとき，アルゴリズム 4.3 で (17), (18) の有限回の適用で $\{n'_{i\to a}(x_i)\}$, $\{m'_{a\to i}(x_i)\}$ が一定値に収束し、

収束値について，

$$b'_i(x_i) \propto \prod_{a\in N(i)} m'_{a\to i}(x_i) \tag{19}$$

を最大にする $x_i \in \mathcal{X}_i$ は，$P'_{X_i}(x_i) := \max_{x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_N} P_{X_1\cdots X_N}(x_1,\cdots,x_N)$ を最大にする $x_i \in \mathcal{X}_i$ に等しく，

$$b'_a(\mathbf{x}_a) \propto f_a(\mathbf{x}_a) \prod_{i\in N(a)} n'_{i\to a}(x_i) \tag{20}$$

を最大にする $\mathbf{x}_a \in \mathcal{X}_a$ は，$P'_{\mathbf{X}_a}(\mathbf{x}_a) := \max_{x_j, j\notin \mathcal{N}(a)} P_{X_1\cdots X_N}(x_1,\cdots,x_N)$ を最大にする $\mathbf{x}_a \in \mathcal{X}_a$ に等しい。

定理 4.1 の証明 (1)

$d_{a,i}$: 因子グラフの変数頂点 $i \in \mathcal{N}$ から，因子頂点 $a \in \mathcal{M}(i)$ の方向にあるもっとも遠い因子頂点までに何個の因子頂点があるか (a 自身も含める) の最大値

$\mathcal{M}_{a,i}$: a の方向にある因子頂点の集合 (a 自身も含める)

$\mathcal{N}_{a,i}$: a の方向にある変数頂点の集合

$\mathcal{Y}_{a,i} := \prod_{j \in \mathcal{N}_{a,i}} \mathcal{X}_j$

$d_{a,i}$ 回の更新で

$$m_{a \to i}(x_i) = \sum_{\mathbf{y}_{a,i} \in \mathcal{Y}_{a,i}} \prod_{c \in \mathcal{M}[a,i]} f_c(\mathbf{x}_c) \tag{13}$$

が成立し，$m_{a \to i}(x_i)$ の値はそれ以上は更新されないことを示す

定理 4.1 の証明 (2)

1. $d_{a,i}=1$ であれば，(12) より、$m_{a\to i}(x_i)=f_a(x_i)$ となり、それ以降更新されることはない。また，(13) が成立している。
2. $$d_{a,i}-1=\max_{j\in\mathcal{N}(a)\backslash\{i\}}\max_{c\in\mathcal{M}(j)\backslash\{a\}}d_{c\to j}$$
   回の (7) の適用で，各 $j\in\mathcal{N}(a)\backslash\{i\}$，$c\in\mathcal{M}(j)\backslash\{a\}$ について
   $$m_{c\to j}(x_j)=\sum_{\mathbf{y}_{c,j}\in\mathcal{Y}_{c,j}}\prod_{e\in\mathcal{M}_{c,j}}f_e(\mathbf{x}_e)$$
   が成立し，それ以上更新されないことを仮定する。

定理 4.1 の証明 (3)

3. さらに (7) を適用して (合計で $d_{a,i}$ 回),

$$
\begin{aligned}
m_{a\to i}(x_i) &= \sum_{\mathbf{x}_{a,i}\in\mathcal{X}_{a,i}} f_a(\mathbf{x}_a) \prod_{j\in\mathcal{N}(a)\setminus\{i\}} \prod_{c\in\mathcal{M}(j)\setminus\{a\}} \sum_{\mathbf{y}_{c,j}\in\mathcal{Y}_{c,j}} \prod_{e\in\mathcal{M}_{c,j}} f_e(\mathbf{x}_e) \\
&= \sum_{x_j\in\mathcal{X}_j, j\in\mathcal{N}(a)\setminus\{i\}} f_a(\mathbf{x}_a) \prod_{j\in\mathcal{N}(a)\setminus\{i\}} \prod_{c\in\mathcal{M}(j)\setminus\{a\}} \sum_{\mathbf{y}_{c,j}\in\mathcal{Y}_{c,j}} \prod_{e\in\mathcal{M}_{c,j}} f_e(\mathbf{x}_e) \\
&= \sum_{\mathbf{y}_{a,i}\in\mathcal{Y}_{a,i}} f_a(\mathbf{x}_a) \prod_{j\in\mathcal{N}(a)\setminus\{i\}} \prod_{c\in\mathcal{M}(j)\setminus\{a\}} \prod_{e\in\mathcal{M}_{c,j}} f_e(\mathbf{x}_e) \\
&= \sum_{\mathbf{y}_{a,i}\in\mathcal{Y}_{a,i}} \prod_{c\in\mathcal{M}_{a,i}} f_c(\mathbf{x}_c)
\end{aligned}
$$

が得られ，またそれ以上更新されないことがわかった。

$\max_{i\in\mathcal{N}} \max_{a\in\mathcal{M}(i)} d_{a\to i}$ 回の更新で (13) が成立し，$\{m_{a\to i}(x_i)\}$ の各値はそれ以上は更新されない。

定理 4.1 の証明 (4)

$$
\begin{aligned}
b_i(x_i) &\propto \prod_{a \in \mathcal{M}(i)} m_{a \to i}(x_i) \\
&= \prod_{a \in \mathcal{M}(i)} \sum_{\mathbf{y}_{a,i} \in \mathcal{Y}_{a,i}} \prod_{c \in \mathcal{M}_{a,i}} f_c(\mathbf{x}_c) \\
&= \sum_{x_j \in \mathcal{X}_j, i \neq j} \prod_{a \in \mathcal{M}(i)} \prod_{c \in \mathcal{M}_{a,i}} f_c(\mathbf{x}_c) \\
&= \sum_{x_j \in \mathcal{X}_j, i \neq j} \prod_{a \in \mathcal{M}} f_a(\mathbf{x}_a).
\end{aligned}
$$

定理 4.1 の証明 (5)

$$
\begin{aligned}
b_a(\mathbf{x}_a) &\propto f_a(\mathbf{x}_a) \prod_{i \in \mathcal{N}(a)} n_{i \to a}(x_i) \\
&= f_a(\mathbf{x}_a) \prod_{i \in \mathcal{N}(a)} \prod_{c \in \mathcal{M}(i) \setminus \{a\}} m_{i \to c}(x_j) \\
&= f_a(\mathbf{x}_a) \prod_{i \in \mathcal{N}(a)} \prod_{c \in \mathcal{M}(i) \setminus \{a\}} \sum_{\mathbf{y}_{c,i} \in \mathcal{Y}_{c,i}} \prod_{e \in \mathcal{M}_{c,i}} f_e(\mathbf{x}_e) \\
&= \sum_{x_i \in \mathcal{X}_i, i \notin \mathcal{N}(a)} f_a(\mathbf{x}_a) \prod_{i \in \mathcal{N}(a)} \prod_{c \in \mathcal{M}(i) \setminus \{a\}} \prod_{e \in \mathcal{M}_{c,i}} f_e(\mathbf{x}_e) \\
&= \sum_{x_i \in \mathcal{X}_i, i \notin \mathcal{N}(a)} \prod_{a \in \mathcal{M}} f_a(\mathbf{x}_a).
\end{aligned}
$$