鈴木讓「ベイジアンネットワーク入門」(培風館)
2. グラフィカルモデル
2.3 Markov ネットワーク: 依存モデルの無向グラフによる表現

鈴木讓

大阪大学

2009 年 11 月 5 日

あらまし

1. Markov ネットワーク
2. Markov 毛布
3. Gibbs ポテンシャル

Markov ネットワーク

$G = (\mathbf{U}, \mathbf{E})$
M: 依存モデル

頂点集合 $\mathbf{U}$ = 確率変数の集合 $\mathbf{U}$

$G = (\mathbf{U}, \mathbf{E})$ が M の I-対応

$\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{U}$

$$\langle \mathbf{X} | \mathbf{Z} | \mathbf{Y} \rangle_G \Longrightarrow I(\mathbf{X}, \mathbf{Z}, \mathbf{Y})_M \tag{25}$$

$G = (\mathbf{U}, \mathbf{E})$ が M の D-対応

$\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{U}$

$$\langle \mathbf{X} | \mathbf{Z} | \mathbf{Y} \rangle_G \Longleftarrow I(\mathbf{X}, \mathbf{Z}, \mathbf{Y})_M$$

$G = (\mathbf{U}, \mathbf{E})$ が M の完全対応

$\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{U}$

$$\langle \mathbf{X} | \mathbf{Z} | \mathbf{Y} \rangle_G \Longleftrightarrow I(\mathbf{X}, \mathbf{Z}, \mathbf{Y})_M$$

M は無向グラフ同型 (無向グラフと同型)

G が M の極小 I-対応 (G が M の Markov ネットワーク)

1. $G = (\mathbf{U}, \mathbf{E})$ が M の I-対応
2. 任意の $\mathbf{E}' \subsetneq \mathbf{E}$ に対して，$G' = (\mathbf{U}, \mathbf{E}')$ が M の I-対応ではない

Markov ネットワーク (続)

(例)

X, Y: 独立な確率変数 ($X, Y, X+Y$ のどの 2 つも独立)

M: $\mathbf{U} := \{X, Y, X+Y\}$ の依存モデル

$I(\{X\}, \{Y\}, \{X+Y\})_M, \quad I(\{Y\}, \{X+Y\}, \{X\})_M,$

$I(\{X+Y\}, \{X\}, \{Y\})_M$

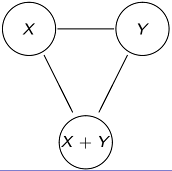
はどれも偽。M の Markov ネットワーク G は，完全グラフ

$$I(\{X\}, \{\}, \{Y\})_M, \quad I(\{Y\}, \{\}, \{X+Y\})_M, \quad I(\{X+Y\}, \{\}, \{X\})_M$$

はどれも真。$\{X, Y\}, \{Y, X+Y\}, \{X+Y, X\} \in \mathbf{E}$ より、

$$\langle \{X\}|\{\}|\{Y\} \rangle_G, \quad \langle \{Y\}|\{\}|\{X+Y\} \rangle_G, \quad \langle \{X+Y\}|\{\}|\{X\} \rangle_G$$

のどれも偽。G は D-対応ではなく，M は無向グラフ同型ではない。

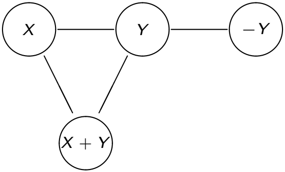


Markov ネットワーク (続)

(例)
X, Y: 確率 1/2 で 0, 1 が生じる独立な確率変数
$\mathbf{U} = \{X, Y, X+Y, -Y\}$
M の Markov ネットワーク

$$\langle \{X\} | \{Y\} | \{-Y\} \rangle_G, \quad I(\{X\}, \{Y\}, \{-Y\})_M$$

$$\langle \{X+Y\} | \{Y\} | \{-Y\} \rangle_G, \quad I(\{X+Y\}, \{Y\}, \{-Y\})_M$$



定理 2.5

M: (5),(6),(9) を満足する依存モデル

定理 2.5

1. 極小 I-対応 $G = (\mathbf{U}, \mathbf{E})$ が存在し、一意
2. G は，完全グラフから，$I(\{X\}, \mathbf{U} - \{X, Y\}, \{Y\})_M$ が真となる $\{X, Y\} \in \mathbf{E}$ を除いて得られる:

$$\{X, Y\} \notin \mathbf{E} \iff I(\{X\}, \mathbf{U} - \{X, Y\}, \{Y\})_M. \qquad (26)$$

証明: テキスト (58-59 ページ)

命題 2.4

$G = (\mathbf{U}, \mathbf{E})$ 無向グラフ
M: 依存モデル

命題 2.4

重複のない $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{U}$ に対して (25) がいえれば，G は M の I-対応。

証明: 命題 2.1、命題 2.1 より

定理 2.6

定理 2.6

依存モデル M が無向グラフ同型
$\Longleftrightarrow I(\cdot,\cdot,\cdot)_M$ が (5),(6),(9) および

$$I(\mathbf{X},\mathbf{Z},\mathbf{Y}) \Longrightarrow I(\mathbf{X},\mathbf{Z}\cup\mathbf{W},\mathbf{Y}), \tag{27}$$

$$I(\mathbf{X},\mathbf{Z},\mathbf{Y}) \Longrightarrow I(\mathbf{X},\mathbf{Z},\{U\}) \vee I(\{X\},\mathbf{Z},\mathbf{Y}),\ \ U \notin \mathbf{Z}. \tag{28}$$

証明: テキスト (59-60 ページ)

Markov 毛布

$\mathbf{Z}$ が $X \in \mathbf{U}$ の Markov 毛布

$$I(\{X\}, \mathbf{Z}, \mathbf{U} - \mathbf{Z} - \{X\})_M, \quad X \notin \mathbf{Z} \tag{29}$$

$\mathbf{Z}$ が $X \in \mathbf{U}$ の Markov 境界

Markov 毛布であって，任意の $\mathbf{Z}' \subsetneq \mathbf{Z}$ に対して

$$I(\{X\}, \mathbf{Z}', \mathbf{U} - \mathbf{Z}' - \{X\})_M, \quad X \notin \mathbf{Z}'$$

が成立しない

$I(\{X\}, \mathbf{Z}, \{\})_M$ のとき、$\mathbf{Z} = \mathbf{U} - \{X\}$ が Markov 毛布
(任意の依存モデル M に対して，Markov 境界が存在)

定理 2.7

$G = (\mathbf{U}, \mathbf{E})$
$\mathbf{B}_G(X)$: $X \in \mathbf{U}$ に隣接する $\mathbf{U}$ の元の集合

定理 2.7

1. 依存モデル M が (5),(6),(9) を満足
$\Longrightarrow$ 各 $X \in \mathbf{U}$ に対して，唯一の Markov 境界 $\mathbf{B}_M(X)$
2. 依存モデル M が (5),(6),(7),(9) を満足
$\Longrightarrow$ 各 $X \in \mathbf{U}$ に対して，$\mathbf{B}_M(X)$ は，(26) の $\mathbf{B}_G(X)$

定理 2.7 (続)

証明: 定理 2.6 の G について,

1. 各 $X \in \mathbf{U}$ で $B_M(X)$ は一意である。実際,

$$I(\{X\}, \mathbf{Z}, \mathbf{U} - \mathbf{Z} - \{X\})_M$$

の各 $\mathbf{Z}$ に (9) を適用すると，それらの共通集合が $B_M(X)$。

2. (7) および (26) より，$B_G(X) \subseteq B_M(X)$。実際

$$\begin{aligned} Y \notin B_M(X) \cup \{X\} &\Longrightarrow I(\{X\}, B_M(X), \{Y\} \cup \mathbf{U} - \{X, Y\} - \mathbf{B}_M(X) \\ &\Longrightarrow I(\{X\}, \mathbf{U} - \{X, Y\}, \{Y\})_M \Longrightarrow Y \notin B_G(X) \end{aligned}$$

3. (26) より、$Y \notin B_G(X) \cup \{X\} \Longleftrightarrow I(\{X\}, \mathbf{U} - \{X, Y\}, \{Y\})_M$
この範囲での $\mathbf{U} - \{X, Y\}$ ($Y \notin \mathbf{U} - \{X\}$) の共通集合が $B_G(X)$。

$$I(\{X\}, B_G(X), \mathbf{U} - B_G(X) - \{X\})_M$$

4. $B_G(X)$ も $B_M(X)$ も Markov 毛布で、$B_M(X)$ が極小かつ一意。

系 2.1

M: $\mu_{\mathbf{U}}(d\mathbf{u}) > 0$, $\mathbf{u} \in \mathbf{U}(\Omega)$ なる確率 $\mu_{\mathbf{U}}$ の依存モデル

系 2.1

1. $X \in \mathbf{U}$ について，$X \notin B_M(X)$.
2. $X, Y \in \mathbf{U}$ について，$Y \in B_M(X) \Longleftrightarrow X \in B_M(Y)$

証明:

1. 定理 2.1 より，(5),(6),(7),(9)。
2. $B_G(X) = B_M(X)$, $X \in \mathbf{U}$
3. 2 条件は，$B_G(X)$, $X \in \mathbf{U}$ の性質。

系 2.2

M: $\mu_{\mathbf{U}}(d\mathbf{u}) > 0$, $\mathbf{u} \in \mathbf{U}(\Omega)$ なる確率 $\mu_{\mathbf{U}}$ の依存モデル
$G = (\mathbf{U}, \mathbf{E})$: M の Markov ネットワーク

系 2.2

$\mathbf{E} = \{\{X, Y\} \mid X \in \mathbf{U}, Y \in \mathbf{B}_M(X)\}$

証明: 定理 2.7 および系 2.1 と同じ条件では，$B_G(X) = B_M(X)$, $X \in \mathbf{U}$

$$\{X, Y\} \in \mathbf{E} \Longleftrightarrow Y \in B_G(X) \Longleftrightarrow Y \in B_M(X)$$

Markov ネットワークの例

例:

$$\Sigma = \{I(\{X\},\{Y\},\{Z\}), I(\{X\},\{Z\},\{W\}), I(\{Y\},\{Z\},\{W\})$$

$$I(\{X,Y\},\{Z\},\{W\}), I(\{X\},\{Y,Z\},\{W\}), I(\{Y\},\{X,Z\},\{W\})\}$$

M: Σ の要素に (5), (12) を適用した $\mathbf{U} = \{X, Y, Z, W\}$ の依存モデル

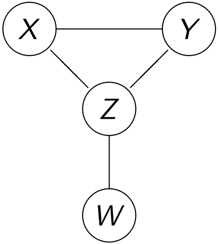
$$I(\{X\},\{Y\},\{Z\})_M, I(\{X\},\{Y,Z\},\{W\})_M, \neg I(\{X\},\{Y\},\{Z,W\})_M$$

1. (8) が満足されない。定理 2.1 より、対応する確率 $\mu_{\mathbf{U}}$ は存在しない。
2. (5), (6), (9) が満足。定理 2.5 より，一意的な極小 I-対応が存在。
3. (7) が満足。定理 2.7 より，(26) に従って，完全グラフから $\{X,W\},\{Y,W\}$ を削除して $B_G(U)$, $U \in \mathbf{U}$ が得られる。

Markov ネットワークの例 (続)

$$B_G(X) = \{Y, Z\}, \quad B_G(Y) = \{X, Z\},$$
$$B_G(Z) = \{X, Y, W\}, \quad B_G(W) = \{Z\}.$$

(29) に従って，Markov 境界を計算してもよい。



系 2.3

M: $\mu_{\mathbf{U}}(d\mathbf{u}) > 0$, $\mathbf{u} \in \mathbf{U}(\Omega)$ なる確率 μ の依存モデル
$G = (\mathbf{U}, \mathbf{E})$

系 2.3

以下の 3 条件は同値。

1. G は M の I-対応である。
2. $\{X, Y\} \notin \mathbf{E} \Longrightarrow I(\{X\}, \mathbf{U} - \{X, Y\}, \{Y\})_M$
3. $X \in \mathbf{U}$ について, $I(\{X\}, \mathbf{B}_G(X), \mathbf{U} - \{X\} - \mathbf{B}_G(X))_M$

証明: 2. $\Rightarrow$ 1. 仮定より, (5),(6),(7),(8),(9) が成立する。定理 2.5 より (26) が極小 I-対応であり、2. であれば, G は I-対応。
1. $\Rightarrow$ 3. $B_G(X)$ の定義から、$\langle\{X\}|B_G(X)|\mathbf{U} - \{X\} - B_G(X)\rangle_G$ が成立。
G が M の I-対応であれば,

$$\langle\{X\}|B_G(X)|\mathbf{U} - \{X\} - B_G(X)\rangle_G \Longrightarrow I(\{X\}, B_G(X), \mathbf{U} - \{X\} - B_G(X))_M$$

3. $\Rightarrow$ 2. 各 $\{X, Y\} \notin \mathbf{E}$ で $\mathbf{W} = \mathbf{U} - \{X, Y\} - B_G(X)$ を (7) に適用

$$I(\{X\}, B_G(X), \mathbf{U} - \{X\} - B_G(X))_M \Longrightarrow I(\{X\}, \mathbf{U} - \{X, Y\}, \{Y\})_M$$

Gibbs ポテンシャル

$\mathbb{U} = \{X_1, \cdots, X_n\}$, $|X_i(\Omega)| < \infty$
$\mathbf{C}_1, \cdots, \mathbf{C}_m \subseteq \mathbf{U}$: $G = (\mathbf{U}, \mathbf{E})$ の極大クリーク
$g_1, \cdots, g_m$: $\mathbf{C}_1(\Omega), \cdots, \mathbf{C}_m(\Omega)$ から非負実数への写像

$$K := \sum_{\mathbf{U}} \prod_{i=1}^{m} g_i(\mathbf{C}_i)$$

Gibbs ポテンシャル

$$P(\mathbf{U}) = K^{-1} \prod_{i=1}^{m} g_i(\mathbf{C}_i) > 0 \tag{30}$$

命題 2.5

命題 2.5

$I(\mathbf{X}, \mathbf{Z}, \mathbf{Y})$
$\Longleftrightarrow$ 各 $\mathbf{x} \in \mathbf{X}(\Omega)$, $\mathbf{y} \in \mathbf{Y}(\Omega)$, $\mathbf{z} \in \mathbf{Z}^*(\Omega) := \{\mathbf{z} \in \mathbf{Z} \mid P_{\mathbf{Z}}(\mathbf{z}) > 0\}$ に対して,

$$P_{\mathbf{XY}|\mathbf{Z}}(\mathbf{x}, \mathbf{y}|\mathbf{z}) = g_{\mathbf{XZ}}(\mathbf{x}, \mathbf{z}) g_{\mathbf{YZ}}(\mathbf{y}, \mathbf{z}) \tag{31}$$

なる $g_{\mathbf{XZ}} : \mathbf{X}(\Omega) \times \mathbf{Z}^*(\Omega) \to [0,1]$, $g_{\mathbf{YZ}} : \mathbf{Y}(\Omega) \times \mathbf{Z}^*(\Omega) \to [0,1]$ が存在。

命題 2.5 (続)

証明：各 $\mathbf{x} \in \mathbf{X}(\Omega), \mathbf{y} \in \mathbf{Y}(\Omega), \mathbf{z} \in \mathbf{Z}^*(\Omega)$ について，(31) を仮定すると、

$$P_{\mathbf{X}|\mathbf{Z}}(\mathbf{x}|\mathbf{z}) = g_{\mathbf{XZ}}(\mathbf{x}, \mathbf{z}) \sum_{\mathbf{y} \in \mathbf{Y}(\Omega)} g_{\mathbf{YZ}}(\mathbf{y}, \mathbf{z}),$$

$$P_{\mathbf{Y}|\mathbf{Z}}(\mathbf{y}|\mathbf{z}) = g_{\mathbf{YZ}}(\mathbf{y}, \mathbf{z}) \sum_{\mathbf{x} \in \mathbf{X}(\Omega)} g_{\mathbf{XZ}}(\mathbf{x}, \mathbf{z}),$$

$$\sum_{\mathbf{x} \in \mathbf{X}(\Omega)} g_{\mathbf{YZ}}(\mathbf{y}, \mathbf{z}) \sum_{\mathbf{y} \in \mathbf{Y}(\Omega)} g_{\mathbf{XZ}}(\mathbf{x}, \mathbf{z}) = 1.$$

3 つの式の辺々をかけて，(1) が得られる。逆に，(1) を仮定すると

$$g_{\mathbf{XZ}}(\mathbf{x}, \mathbf{z}) := P_{\mathbf{X}|\mathbf{Z}}(\mathbf{x}|\mathbf{z}),$$

$$g_{\mathbf{YZ}}(\mathbf{y}, \mathbf{z}) := P_{\mathbf{Y}|\mathbf{Z}}(\mathbf{y}|\mathbf{z})$$

とおけば，(31) が成立。

定理 2.8

定理 2.8

P が G の Gibbs ポテンシャルであるとき、G は P の I-対応

証明: $J(X) := \{j \mid X \in \mathbf{C}_j\}$, $B_G(X) \cup \{X\} = \bigcup_{j \in J(X)} \mathbf{C}_j$
(30) より，各 $X \in \mathbf{U}$ で

$$P(\mathbf{U}) = \prod_{j \in J(X)} g_j(\mathbf{C}_j) \cdot \prod_{j \notin J(X)} g_j(\mathbf{C}_j) = h_1(B_G(X) \cup \{X\}) h_2(\mathbf{U} - \{X\}),$$

$$P(B_G(X)) = h_3(B_G(X) \cup \{X\})$$

なる関数 h_1, h_2, h_3 が存在。したがって，

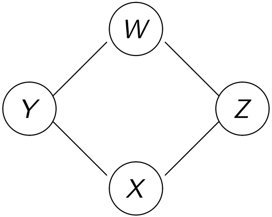
$$P(\mathbf{U}|B_G(X)) = \frac{h_1(B_G(X) \cup \{X\})}{h_3(B_G(X) \cup \{X\})} \cdot h_2(\mathbf{U} - \{X\}).$$

命題 2.5 より，$I(\{X\}, B_G(X), \mathbf{U} - \{X\} - B_G(X))$，$X \in \mathbf{U}$ が成立。
系 2.3 より G は P の I-対応。

Gibbs ポテンシャルの例

$\mathbf{U} := \{X, Y, Z, W\}$
$G = (\mathbf{U}, \mathbf{E})$
$\mathbf{C}_1 := \{X, Y\}$, $\mathbf{C}_2 := \{X, Z\}$, $\mathbf{C}_3 := \{Y, W\}$, $\mathbf{C}_4 := \{Z, W\}$
Gibbs ポテンシャル

$$P_{XYZW}(x, y, z, w) \propto g_1(x, y) g_2(x, z) g_3(y, w) g_4(z, w)$$



Gibbs ポテンシャルの例 (続)

$$
\begin{aligned}
& P_{W|XYZ}(w|x,y,z) \\
= \ & \frac{P_{XYZW}(x,y,z,w)}{\sum_w P_{XYZW}(x,y,z,w)} \\
= \ & \frac{g_1(x,y)g_2(x,z)g_3(y,w)g_4(z,w)}{g_1(x,y)g_2(x,z)\sum_w g_3(y,w)g_4(z,w)} \\
= \ & \frac{\sum_x g_1(x,y)g_2(x,z)\cdot g_3(y,w)g_4(z,w)}{\sum_x g_1(x,y)g_2(x,z)\cdot \sum_w g_3(y,w)g_4(z,w)} \\
= \ & \frac{\sum_x g_1(x,y)g_2(x,z)g_3(y,w)g_4(z,w)}{\sum_{x,w} g_1(x,y)g_2(x,z)g_3(y,w)g_4(z,w)} \\
= \ & \frac{P_{YZW}(y,z,w)}{P_{YZ}(y,z)} \\
= \ & P_{W|YZ}(w|y,z).
\end{aligned}
$$

Gibbs ポテンシャルの例 (続)

同様に、$P_{Y|XZW}(y|x,z,w) = P_{Y|XW}(y|x,w)$

$$I(\{X\},\{Y,Z\},\{W\})_P, \quad I(\{Y\},\{X,W\},\{Z\})_P$$

G が意味しているすべての条件付独立性が成立 (I-対応)。

$$\neg I(\{X\},\{\},\{W\})_P, \quad \neg I(\{Y\},\{\},\{Z\})_P$$

$$\neg\langle\{X\}|\{\}|\{W\}\rangle_G, \quad \neg\langle\{Y\}|\{\}|\{Z\}\rangle_G$$

(D-対応)

P できまる依存モデル M は、無向グラフ同型

定理 2.9

定理 2.9

P の I-対応となる G で定理 2.2 の 4 条件が成立するとき、P は G の極大クリークにおける確率の積をその辺の確率の積で割った形でかける。

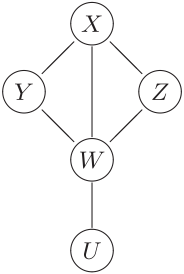
(例)

$$\begin{aligned} P(X,Y,Z,W,U) &= P(X,Y,W)P(Z|X,W)P(U|W) \\ &= \frac{P(X,Y,W)P(X,Z,W)P(W,U)}{P(X,W)P(W)}, \end{aligned}$$

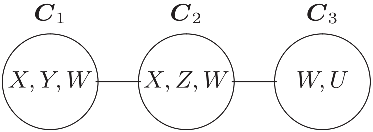
$$\begin{aligned} P(X,Y,Z,W,U) &= P(X,Z,W)P(Y|X,W)P(U|W) \\ &= \frac{P(X,Z,W)P(X,Y,W)P(W,U)}{P(X,W)P(W)} \end{aligned}$$

定理 2.9 (続)

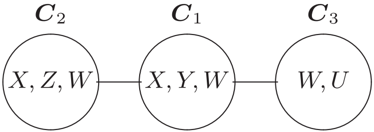
(a)



(b)



(c)



$\{X, Y, W\} \cap \{W, U\} \subseteq \{X, Z, W\}$

$\{X, Z, W\} \cap \{W, U\} \subseteq \{X, Y, W\}$

定理 2.9 (続)

証明: 定理 2.2 より接合木が存在する。$\mathbf{C}_1, \cdots, \mathbf{C}_m$ をそのクリークとする。G が P の I-対応であることと (11) より，接合木 T は P の I-対応:

$$\begin{aligned}\langle \mathbf{C}_i | \mathbf{C}_j | \mathbf{C}_k \rangle_G &\implies \langle \mathbf{C}_i - \mathbf{C}_j | \mathbf{C}_j | \mathbf{C}_k - \mathbf{C}_j \rangle_G \\ &\implies I(\mathbf{C}_i - \mathbf{C}_j, \mathbf{C}_j, \mathbf{C}_k - \mathbf{C}_j) \\ &\implies I(\mathbf{C}_i, \mathbf{C}_j, \mathbf{C}_k),\end{aligned}$$

$$P(\mathbf{U}) = \prod_{i=1}^{m} P(\mathbf{C}_i | \mathbf{C}_1, \cdots, \mathbf{C}_{i-1}) = \prod_{i=1}^{m} P(\mathbf{C}_i | \mathbf{C}_{j(i)})$$

なる $\mathbf{C}_i$ の添え字 i のつけ方，および $j(i) \in \{1, \cdots, m\}$ を設定できる。G が P の I-対応であることと，$\langle \mathbf{C}_i | \mathbf{C}_i \cap \mathbf{C}_{j(i)} | \mathbf{C}_{j(i)} - \mathbf{C}_i \rangle$ より，

$$P(\mathbf{C}_i | \mathbf{C}_{j(i)}) = P(\mathbf{C}_i | \mathbf{C}_i \cap \mathbf{C}_{j(i)})$$

$$P(\mathbf{U}) = \prod_{i=1}^{m} P(\mathbf{C}_i | \mathbf{C}_i \cap \mathbf{C}_{j(i)}) = \prod_{i=1}^{m} \frac{P(\mathbf{C}_i)}{P(\mathbf{C}_i \cap \mathbf{C}_{j(i)})}$$