

鈴木讓「ベイジアンネットワーク入門」(培風館)

4. 確率的推論

4.3 統計力学との関連

鈴木讓

大阪大学

2009年1月28日(木)

# あらまし

- 1 統計力学
- 2 領域
- 3 領域グラフ
- 4 接続グラフ
- 5 クラスタ変分法
- 6 一般化ビリーフ・プロパゲーション

## Boltzmann の法則

$E(\mathbf{x})$ : 粒子  $X_1, \dots, X_N$  の状態に対応するエネルギー

$T$ : 絶対温度

$$p(\mathbf{x}) := P_{X_1 \dots X_N}(x_1, \dots, x_N) = \frac{1}{Z(T)} e^{-E(\mathbf{x})/T} \quad (22)$$

$$Z := Z(1) = \sum_{\mathbf{x}} e^{-E(\mathbf{x})}$$

$$E(\mathbf{x}) := - \sum_{a \in \mathcal{M}} \log f_a(\mathbf{x}_a)$$

Helmholts エネルギー

$$F_H := - \log Z$$

## 統計力学 (続)

$$0 \leq b(\mathbf{x}) \leq 1, \quad \sum_{\mathbf{x}} b(\mathbf{x}) = 1$$

平均エネルギー:

$$U(b) := \sum_{\mathbf{x}} b(\mathbf{x}) E(\mathbf{x}), \quad (23)$$

エントロピー:

$$H(b) := - \sum_{\mathbf{x}} b(\mathbf{x}) \log b(\mathbf{x}) \quad (24)$$

自由エネルギー:

$$F(b) := U(b) - H(b) \quad (25)$$

$$D(b||p) := \sum_{\mathbf{x}} b(\mathbf{x}) \log \frac{b(\mathbf{x})}{p(\mathbf{x})} \geq 0$$

$F_H$  の上界である  $F(b) = F_H + D(b||p)$  の最小化をめざす

## 例 4.11 平均場近似

$$b(\mathbf{x}) := \prod_{i=1}^N b_i(x_i) \quad (26)$$

$$U(\{b_1, \dots, b_N\}) := - \sum_{a \in \mathcal{M}} \sum_{\mathbf{x}_a} \prod_{i \in \mathcal{N}(a)} b_i(x_i) \log f_a(\mathbf{x}_a),$$

$$H(\{b_1, \dots, b_N\}) := - \sum_{i=1}^N \sum_{x_i} b_i(x_i) \log b_i(x_i).$$

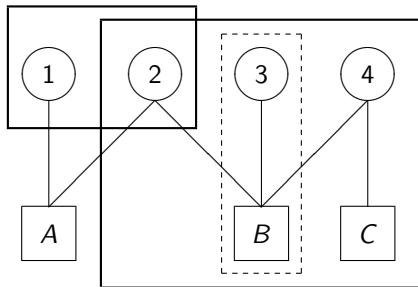
$F(\{b_1, \dots, b_N\})$  に Lagrange の未定乗数法を適用  
計算量が少なく、任意の  $S \subseteq \mathcal{N}$  について

$$b(\mathbf{x}_S) = \sum_{\mathbf{x}_{\mathcal{N} \setminus S}} b(\mathbf{x}_{\mathcal{N} \setminus S})$$

## 領域

## 領域

$a \in M \cap R \implies \mathcal{N}(a) \cup \{a\} \in R$  なる  $\mathcal{N} \cup M$  の部分集合



□ 領域、⋯ 領域でない

## 領域集合

$\mathcal{R}$ : 領域集合

領域定数  $\{c_R\}_{R \in \mathcal{R}}$

$$\sum_{R \in \mathcal{R}} c_R I[i \in R] = 1, \quad i \in \mathcal{N}, \quad (27)$$

$$\sum_{R \in \mathcal{R}} c_R I[a \in R] = 1, \quad a \in \mathcal{M} \quad (28)$$

$$I[i \in R] = \begin{cases} 1, & i \in R \\ 0, & i \notin R \end{cases}, \quad I[a \in R] = \begin{cases} 1, & a \in R \\ 0, & a \notin R \end{cases}$$

領域集合  $\mathcal{R}$  が妥当

領域定数  $\{c_R\}_{R \in \mathcal{R}}$  が存在

## 領域に基づく自由エネルギー

領域に基づく平均エネルギー:

$$U_{\mathcal{R}}(\{b_{\mathcal{R}}\}) := \sum_{R \in \mathcal{R}} c_R U_R(b_R), \quad (29)$$

領域に基づくエントロピー:

$$H_{\mathcal{R}}(\{b_{\mathcal{R}}\}) := \sum_{R \in \mathcal{R}} c_R H_R(b_R) \quad (30)$$

領域に基づく自由エネルギー:

$$F_{\mathcal{R}}(\{b_{\mathcal{R}}\}) := U_{\mathcal{R}}(\{b_{\mathcal{R}}\}) - H_{\mathcal{R}}(\{b_{\mathcal{R}}\}) \quad (31)$$



## 領域に基づく自由エネルギーの最小化

$$0 \leq b_R(\mathbf{x}_R) \leq 1 \quad (32)$$

$$\sum_{\mathbf{x}_R \in \mathcal{X}_R} b_R(\mathbf{x}_R) = 1 \quad (33)$$

$$U_R(b_R) := \sum_{\mathbf{x}_R \in \mathcal{X}_R} b(\mathbf{x}_R) E(\mathbf{x}_R)$$

$$H_R(b_R) := - \sum_{\mathbf{x}_R \in \mathcal{X}_R} b(\mathbf{x}_R) \log b(\mathbf{x}_R)$$

領域集合  $\mathcal{R}$  が妥当のとき

- $U_{\mathcal{R}}(\{b_{\mathcal{R}}\}) = U(b)$
- $H_{\mathcal{R}}(\{b_{\mathcal{R}}\}) \neq H(b)$

## 例 4.13

$$\mathcal{R} = \{R\}, R = \mathcal{N} \cup \mathcal{M}$$

$c_R = 1$  となり、(29)–(31) は (23)–(25) に一致

## 例 4.14 Bethe 近似

$$\mathcal{R} = \mathcal{R}_L \cup \mathcal{R}_S,$$

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

$$\mathcal{R}_S = \{\{i\} \mid i \in \mathcal{N}\}$$

$$c_R = \begin{cases} 1, & R \in \mathcal{R}_L \\ 1 - d_i, & R \in \mathcal{R}_S \end{cases}$$

$d_i$ :  $\{i\} = R \in \mathcal{R}_S$  なる  $i \in \mathcal{N}$  に連結している因子頂点の個数

$$F_{\text{Bethe}} := U_{\mathcal{R}}(\{b_R\}) - H_{\mathcal{R}}(\{b_R\})$$

$$U_{\mathcal{R}}(\{b_R\}) = - \sum_{a \in \mathcal{M}} \sum_{\mathbf{x}_a \in \mathcal{X}_a} b(\mathbf{x}_a) \log f_a(\mathbf{x}_a), \quad (34)$$

$$H_{\mathcal{R}}(\{b_R\}) = - \sum_{a \in \mathcal{M}} \sum_{\mathbf{x}_a \in \mathcal{X}_a} b_a(\mathbf{x}_a) \log b_a(\mathbf{x}_a) + \sum_{i \in \mathcal{N}} (d_i - 1) \sum_{x_i \in \mathcal{X}_i} b_i(x_i) \log b_i(x_i)$$

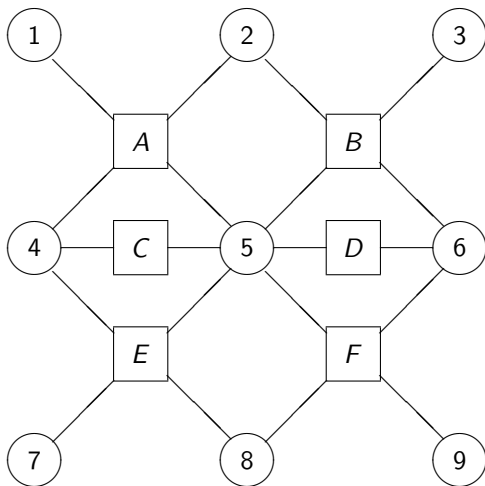
## 例 4.14 Bethe 近似 (続)

$$\begin{aligned}0 &\leq b_i(x_i) \leq 1, \quad i \in \mathcal{N}, \\0 &\leq b_a(\mathbf{x}_a) \leq 1, \quad a \in \mathcal{M}.\end{aligned}$$

$$p(\mathbf{x}) = P_{X_1 \dots X_N}(x_1, \dots, x_N) = \frac{\prod_{a \in \mathcal{M}} P_{\mathbf{x}_a}(\mathbf{x}_a)}{\prod_{i \in \mathcal{N}} P_{X_i}(x_i)^{d_i-1}} \quad (35)$$

- $b(\mathbf{x}) = p(\mathbf{x})$  のとき、 $H_{\mathcal{R}}(\{b_{\mathcal{R}}\}) = H(b)$ 、 $F_H = F_{Bethe}$
- (35) は、因子グラフが巡回経路を含まないことと同値

## 例 4.14 Bethe 近似 (続)



## 例 4.14 Bethe 近似 (続)

$R \in \mathcal{R}_L$	$C_R$
$R_A = \{A, 1, 2, 4, 5\}$	$C_A = 1$
$R_B = \{B, 2, 3, 5, 6\}$	$C_B = 1$
$R_C = \{C, 4, 5\}$	$C_C = 1$
$R_D = \{D, 5, 6\}$	$C_D = 1$
$R_E = \{E, 4, 5, 7, 8\}$	$C_E = 1$
$R_F = \{F, 5, 6, 8, 9\}$	$C_F = 1$

$R \in \mathcal{R}_S$	$C_R$
$R_1 = \{1\}$	$C_1 = 1 - 1 = 0$
$R_2 = \{2\}$	$C_2 = 1 - 2 = -1$
$R_3 = \{3\}$	$C_3 = 1 - 1 = 0$
$R_4 = \{4\}$	$C_4 = 1 - 3 = -2$
$R_5 = \{5\}$	$C_5 = 1 - 6 = -5$
$R_6 = \{6\}$	$C_6 = 1 - 3 = -2$
$R_7 = \{7\}$	$C_7 = 1 - 1 = 0$
$R_8 = \{8\}$	$C_8 = 1 - 2 = -1$
$R_9 = \{9\}$	$C_9 = 1 - 1 = 0$

## 定理 4.4

## 定理 4.4

- 定理 4.1 で計算される  $b_a(\mathbf{x}_a) > 0$ ,  $b_i(x_i) > 0$  なる解  $\{b_a(\mathbf{x}_a)\}_{a \in \mathcal{M}}$ ,  $\{b_i(x_i)\}_{i \in \mathcal{N}}$
- Bethe 自由エネルギー  $F_{Bethe}$  を極小とする  $\hat{b}_a(\mathbf{x}_a) > 0$ ,  $\hat{b}_i(x_i) > 0$  なる解  $\{\hat{b}_a(\mathbf{x}_a)\}_{a \in \mathcal{M}}$ ,  $\{\hat{b}_i(x_i)\}_{i \in \mathcal{N}}$

証明のスケッチ:  $F_{Bethe}$  に Lagrange の未定乗数法を適用  
未定乗数および  $\{b_a(\mathbf{x}_a)\}_{a \in \mathcal{M}}$ ,  $\{b_i(x_i)\}_{i \in \mathcal{N}}$  で偏微分して 0 とおく。

## 領域グラフ

$\mathcal{R}$ : 領域集合

$D = (\mathbf{V}, \vec{\mathbf{E}})$ : 領域  $R, R' \in \mathcal{R}$  について,  $\mathbf{V} = \mathcal{R}$ ,

$$(R, R') \in \vec{\mathbf{E}} \implies R \supsetneq R'$$

$A(R)$ :  $R \in \mathcal{R}$  の先祖全体

$D$  が  $\mathcal{R}$  の領域グラフ

- ①  $c_R = 1 - \sum_{R' \in A(R)} c_{R'}$ .
- ② 各  $i \in I := \mathcal{N} \cup \mathcal{M}$  を含む領域  $R$  が連結されている。

なる領域定数  $\{c_R\}_{R \in \mathcal{R}}$  が存在

## 命題 4.6

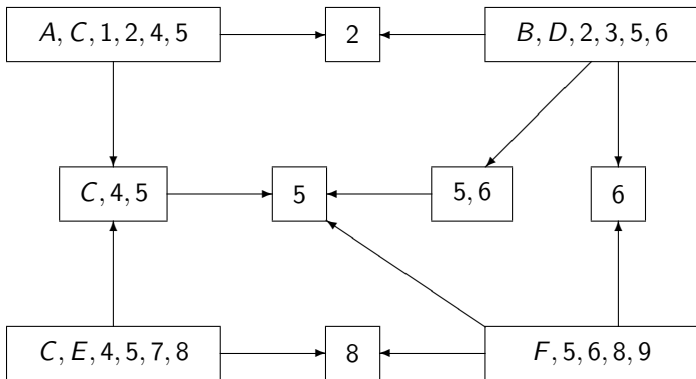
領域グラフが存在  $\implies$  領域集合は妥当 (領域定数が存在)



## 例 4.15

$R \in \mathcal{R}$	$A(R)$ の要素	$C_R$
$R_{AC1245}$ $= \{A, C, 1, 2, 4, 5\}$		$C_{AC1245} = 1$
$R_{BD2356}$ $= \{B, D, 2, 3, 5, 6\}$		$C_{BD2356} = 1$
$R_{CE4578}$ $= \{C, E, 4, 5, 7, 8\}$		$C_{CE4578} = 1$
$R_{F5689}$ $= \{F, 5, 6, 8, 9\}$		$C_{F5689} = 1$
$R_{C45} = \{C, 4, 5\}$	$R_{AC1245}, R_{CE4578}$	$C_{C45} = 1 - 2 \cdot 1 = -1$
$R_2 = \{2\}$	$R_{AC1245}, R_{BD2356}$	$C_2 = 1 - 2 \cdot 1 = -1$
$R_6 = \{6\}$	$R_{BD2356}, R_{F5689}$	$C_6 = 1 - 2 \cdot 1 = -1$
$R_8 = \{8\}$	$R_{CE4578}, R_{F5689}$	$C_8 = 1 - 2 \cdot 1 = -1$
$R_{56} = \{5, 6\}$	$R_{BD2356}$	$C_{56} = 1 - 1 \cdot 1 = 0$
$R_5 = \{5\}$	$R_{AC1245}, R_{BD2356}, R_{CE4578},$ $R_{F5689}, R_{C45}, R_{56}$	$C_5 = 1 - 4 \cdot 1 - 1 \cdot 0$ $- 1 \cdot (-1) = -2$

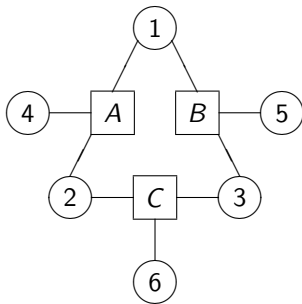
## 例 4.15 (続)



## 例 4.16: 命題 4.6 の逆は成立しない

$$\mathcal{R} = \{\{A, 1, 2, 4\}, \{B, 1, 3, 5\}, \{C, 2, 3, 6\}, \{1, 2, 3\}\}$$

- $\mathcal{R}$  は妥当:  $1, 1, 1, -1$  は (27), (28) を満足
- $\mathcal{R}$  の領域グラフが存在しない:  $\{1, 2, 3\}$  は  $\{A, 1, 2, 4\}, \{B, 1, 3, 5\}, \{C, 2, 3, 6\}$  のいずれの部分集合でもない

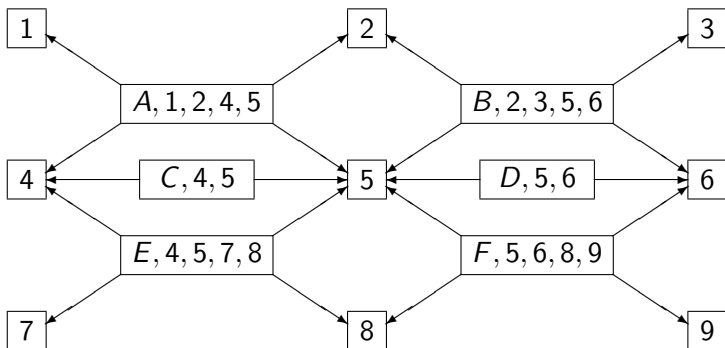


## 例 4.17: Bethe 近似を適用すれば、領域グラフが存在

$R \in \mathcal{R}_L$  から  $i \in R$  なる  $\{i\} \in \mathcal{R}_S$  に有向辺

- 世代が高々2つ (有向経路の長さが高々1つ)
- $c_R = 1 - d_R$ ,  $R \in \mathcal{R}_S$

( $d_R$ :  $\mathcal{R}_L$  から入ってくる有向辺の個数)



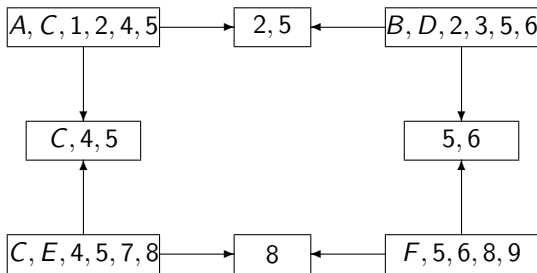
## 例 4.18: 接続グラフ

$\mathcal{R}_S$  に複数の要素を入れても、必ず領域グラフが得られる

- 世代が高々2つ (有向経路の長さが高々1つ)
- $c_R = 1 - d_R, R \in \mathcal{R}_S$

$$\mathcal{R}_L = \{\{A, C, 1, 2, 4, 5\}, \{B, D, 2, 3, 5, 6\}, \{C, E, 4, 5, 7, 8\}, \{F, 5, 6, 8, 9\}\}$$

$$\mathcal{R}_S = \{\{2, 5\}, \{C, 4, 5\}, \{5, 6\}, \{8\}\}$$



## 接続木は、接続グラフの特殊ケース

- ①  $d_R = 2$  ( $\iff c_R = -1$ ),  $R \in \mathcal{R}_S$   
 $R \in \mathcal{R}_S$  が接続木の辺
- ②  $R \in \mathcal{R}_S$  は,  $a \in \mathcal{M}$  を含まない  
 $\mathcal{M}$  の部分集合が, 接続木の頂点としての役割を果たす

## Bethe 近似

接続グラフで表現できても, 接続木では表現できない

## クラスター変分法 (Kikuchi 近似)

$$\mathcal{R} = \mathcal{R}_1 \cup \dots \cup \mathcal{R}_K:$$

- ①  $I := \mathcal{N} \cup M$  の各要素が 1 つ以上の領域  $R$  で現れるような  $\mathcal{R}_0$  を選択 (他の要素の真部分集合になっているものは削除)
- ②  $R, R' \in \mathcal{R}_0$  ( $R \neq R'$ ) で、 $R \cap R' \in \mathcal{R}_1$  となる  $\mathcal{R}_1$  を選択 (他の要素の真部分集合になっているものは削除)
- ③  $R, R' \in \mathcal{R}_0 \cup \mathcal{R}_1$  ( $R \neq R'$ ) で、 $R \cap R' \in \mathcal{R}_2$  となるように  $\mathcal{R}_2$  を選択 (すでに  $\mathcal{R}_1$  に現れているもの、 $\mathcal{R}_2$  の他の要素の真部分集合になっているものは削除)
- ④ これを繰り返して、 $\mathcal{R}_3, \dots, \mathcal{R}_K$  を得る。

## Kikuchi 近似

世代数の上限が 2 を超える場合がある

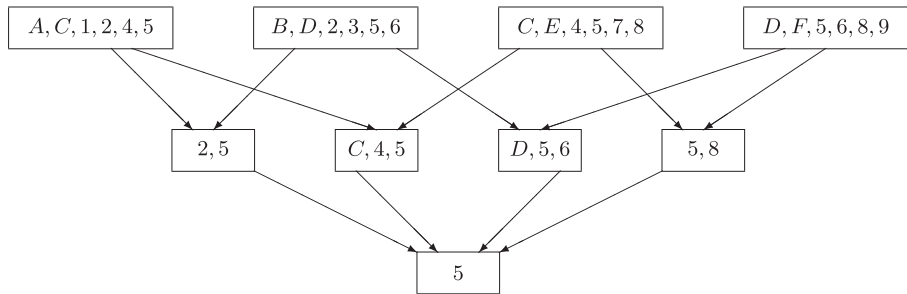
# 例 4.19: クラスタ変分法 (Kikuchi 近似)

$$\mathcal{R}_0 = \{\{A, C, 1, 2, 4, 5\}, \{B, D, 2, 3, 5, 6\}, \{C, E, 4, 5, 7, 8\}, \{F, 5, 6, 8, 9\}\}$$

$$c_R = 1$$

$$\mathcal{R}_1 = \{\{2, 5\}, \{C, 4, 5\}, \{D, 5, 6\}, \{5, 8\}\}, \quad c_R = 1 - 2 = -1$$

$$\mathcal{R}_2 = \{\{5\}\}, \quad c_R = 1 - 4 \cdot 1 - 4 \cdot (-1) = 1$$





## 命題 4.7

## 命題 4.7

領域グラフが存在し、因子グラフが巡回経路をもたないとき、自由エネルギーは領域に基づく自由エネルギーに一致

$$c_{A134} = c_{C346} = c_{B245} = c_{D457} = 1, \quad c_{34} = c_{45} = -1, \quad c_4 = -1$$

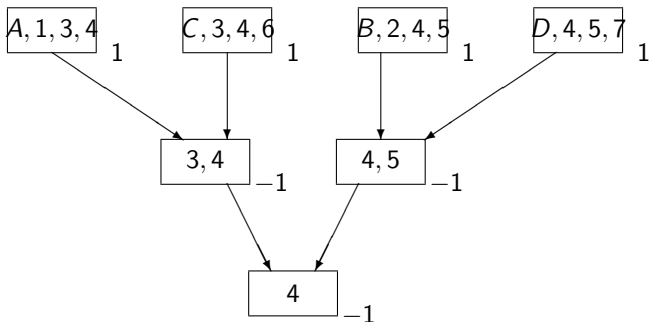
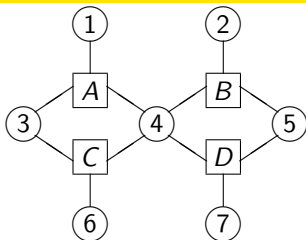
$$p(x_1, x_3, x_4, x_6) = \frac{p_A(x_1, x_3, x_4)p_C(x_3, x_4, x_6)}{p(x_3, x_4)}$$

$$p(x_2, x_4, x_5, x_7) = \frac{p_B(x_2, x_4, x_5)p_D(x_4, x_5, x_7)}{p(x_4, x_5)}$$

$$p(x_1, \dots, x_7) = \frac{p(x_1, x_3, x_4, x_6)p(x_2, x_4, x_5, x_7)}{p(x_4)}$$

$$p(x_1, \dots, x_7) = p_A(x_1, x_3, x_4)p_C(x_3, x_4, x_6)p_B(x_2, x_4, x_5)p_D(x_4, x_5, x_7) \cdot p(x_4)^{-1}p(x_3, x_4)^{-1}p(x_4, x_5)^{-1} = \prod_{R \in \mathcal{R}} p_R(\mathbf{x}_R)^{c_R}$$

# 例 4.20



## ビリーフ・プロパゲーションの一般化

$p_R$ :  $R \in \mathcal{R}$  の親の頂点の個数

$$q_R := \begin{cases} 1 & p_R = 0 \\ (1 - c_R)/p_R & p_R \neq 0 \end{cases}$$

$c_R \neq 1 - 2p_R$  を仮定

$c_R = 1 - 2p_R$  なら、領域定数  $\{c_R\}_{R \in \mathcal{R}}$  が異なる領域グラフが存在

親  $R'$  の先祖の頂点  $R''$  が存在 (存在しなければ,  $c_R = 1 - p_R$ )

$$(R'', R) \in \vec{\mathbf{E}} \implies (R', R) \in \vec{\mathbf{E}}$$

$$(R', R) \in \vec{\mathbf{E}} \implies (R'', R) \in \vec{\mathbf{E}}$$

$$\beta_R := (2 - q_R)^{-1}$$

(Bethe 近似なら、 $q_R = \beta_R = 1, R \in \mathcal{R}$ )

## ビリーフ・プロパゲーションの一般化 (続)

 $R \in \mathcal{R}, P \in \mathcal{P}(R), C \in \mathcal{C}(R)$ 

$$n_{R \rightarrow P}(\mathbf{x}_R) = F_R(\mathbf{x}_R) \prod_{P' \in \mathcal{P}(R) \setminus \{P\}} M_{P' \rightarrow R}(\mathbf{x}_R) \prod_{C \in \mathcal{C}(R)} N_{C \rightarrow R}(\mathbf{x}_C) \quad (36)$$

$$m_{R \rightarrow C}(\mathbf{x}_C) = \sum_{\mathbf{x}_R \setminus \mathbf{x}_C} F_R(\mathbf{x}_R) \prod_{P \in \mathcal{P}(R)} M_{P \rightarrow R}(\mathbf{x}_R) \prod_{C' \in \mathcal{C}(R) \setminus \{C\}} N_{C' \rightarrow R}(\mathbf{x}_{C'}) \quad (37)$$

$$b_R(\mathbf{x}_R) = F_R(\mathbf{x}_R) \prod_{C \in \mathcal{C}(R)} N_{C \rightarrow R}(\mathbf{x}_C) \prod_{P \in \mathcal{P}(R)} M_{P \rightarrow R}(\mathbf{x}_P) \quad (38)$$

$$F_R(\mathbf{x}_R) := \left( \prod_{a \in R \cap \mathcal{M}} f_a(\mathbf{x}_a) \right)^{c_R}$$

$$N_{R \rightarrow P}(\mathbf{x}_R) := n_{R \rightarrow P}(\mathbf{x}_R)^{\beta_R} m_{P \rightarrow R}(\mathbf{x}_R)^{\beta_R - 1} \quad (39)$$

$$M_{R \rightarrow C}(\mathbf{x}_C) := n_{C \rightarrow R}(\mathbf{x}_C)^{\beta_C - 1} m_{R \rightarrow C}(\mathbf{x}_C)^{\beta_C} \quad (40)$$

Bethe 近似 ( $\beta_R = 1$ ) の場合

$N_{R \rightarrow P}(\mathbf{x}_R)$ :  $R \in \mathcal{R}$  から  $P \in \mathcal{P}(R)$  (有向辺と逆の方向)

$M_{R \rightarrow C}(\mathbf{x}_R)$ :  $R \in \mathcal{R}$  から  $C \in \mathcal{C}(R)$  (有向辺と同じ方向)

$$F_R(\mathbf{x}_R) = \begin{cases} f_a(\mathbf{x}_a), & R \in \mathcal{R}_L \\ 1, & R \in \mathcal{R}_S, \end{cases}$$

$$N_{R \rightarrow P}(\mathbf{x}_R) = n_{R \rightarrow P}(\mathbf{x}_R), \quad R \in \mathcal{R}_S, \quad P \in \mathcal{P}(R) = \mathcal{R}_L,$$

$$M_{R \rightarrow C}(\mathbf{x}_R) = m_{R \rightarrow C}(\mathbf{x}_R), \quad R \in \mathcal{R}_L, \quad C \in \mathcal{C}(R) = \mathcal{R}_S,$$

$$b_R(\mathbf{x}_R) = \begin{cases} b_a(\mathbf{x}_a), & R = \{a\} \cup \mathcal{N}(a) \in \mathcal{R}_L \\ b_i(\mathbf{x}_i), & R = \{i\} \in \mathcal{R}_S \end{cases}$$

## 定理 4.5

## 定理 4.5: 一般化ビリーフプロパゲーション

- (36), (37), (39), (40) の更新で  $b_R(\mathbf{x}_R) > 0$  の (38) の解  $\{b_R(\mathbf{x}_R)\}$
- (33) および  $R \in \mathcal{R}$ ,  $C \in \mathcal{C}(R)$  で  $b_C(\mathbf{x}_C) = \sum_{\mathbf{x}_{R \setminus C}} b_R(\mathbf{x}_R)$  で、  
 $F_{\mathcal{R}}(\{b_R\})$  を極小とする  $\hat{b}_R(\mathbf{x}_R) > 0$  なる解  $\{\hat{b}_R(\mathbf{x}_R)\}_{R \in \mathcal{R}}$

接続木や一般の領域グラフに変換して、ビリーフプロパゲーションを適用

- 因子グラフが巡回経路をもつ場合でも、巡回経路をもたない領域グラフが得られれば、 $\{b_R(\mathbf{x})\}_{R \in \mathcal{R}}$  が正確に計算できる
- 領域の頂点の数の最大値に対して、指数的な計算がかかる
- 領域の頂点の数の最大値を計算することは、 $\mathcal{NP}$  困難