

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

4.2 確率的推論における NP 困難性

鈴木讓

大阪大学

2009 年 1 月 21 日 (木)

あらまし

- ① \mathcal{P} と \mathcal{NP}
- ② 確率的推論と最大事後確率設定 \mathcal{NP} の困難性

問題の実行時間

計算機: 有限の 2 進列の入力に対して

- 停止して、有限の 2 進列を出力
- 停止しない

アルゴリズム

すべての入力に対して停止する対応 $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$ の記述

アルゴリズム A の実行時間 T_A

A への入力 $\sigma \in \{0, 1\}^*$ から実行時間 $T_A(\sigma)$ への対応

入力長が $n := |\sigma|$ のときの実行時間 T_n

$$T_n := \min_A \max_{\sigma \in \{0,1\}^n} T_A(\sigma)$$

クラス P

多項式時間で解ける

有限個の n を除いて、 $T_n \leq n^k$ となる k が存在

- 計算機が異なれば T_n は異なる
- 多項式で解けるか否かは、決定性 TM であれば同じ

決定問題 $f : \{0, 1\}^* \rightarrow \{0, 1\}$

入力は

- ① “1” が出力される 2 進列の集合 Σ
- ② “0” が出力される 2 進列の集合 $\{0, 1\}^* \setminus \Sigma$

に大別できる (言語 $\Sigma \subseteq \{0, 1\}^*$ の認識)

P

決定性 TM で、多項式時間で解ける決定問題の集合

クラス NP

決定性 TM 命令の実行が毎回 1 個のみ実行

非決定性 TM 命令の実行が毎回複数同時に実行可 (仮想的)
1 個の決定性 TM で Σ の認識が完了すればよい

NP

非決定性 TM で、多項式時間で解ける決定問題の集合

- $P \subseteq NP$
- $P \neq NP$ であることが強く予想されている

NP 完全

言語 Σ は言語 Σ' に多項式時間で還元 ($\Sigma \propto \Sigma'$)

$f : \{0, 1\}^* \rightarrow \{0, 1\}^*$ s.t.

- $f(\Sigma) \subseteq \Sigma'$
- $f(\{0, 1\}^* \setminus \Sigma) \subseteq \{0, 1\}^* \setminus \Sigma'$

を多項式時間で解く決定性 TM のアルゴリズムが存在

NP 完全な言語

任意の $\Sigma \in \text{NP}$ に対して $\Sigma \propto \Sigma'$ となる言語 Σ'

$\mathcal{P} \neq \mathcal{NP}$ を仮定すると

命題 4.2

- ① $\Sigma \propto \Sigma', \Sigma' \propto \Sigma'' \implies \Sigma \propto \Sigma''$
- ② $\Sigma \propto \Sigma', \Sigma' \in \mathcal{P} \implies \Sigma \in \mathcal{P}$
- ③ Σ が \mathcal{NP} 完全, $\Sigma \propto \Sigma' \implies \Sigma'$ が \mathcal{NP} 完全

$\Sigma \in \mathcal{NP}$ 完全を示せば、 $\Sigma \notin \mathcal{P}$ の強い証拠になる

$\Sigma \in \mathcal{NP}$ 完全, $\Sigma \in \mathcal{P}$ なる Σ が存在 $\implies \mathcal{P} = \mathcal{NP}$ (予想と矛盾)

充足可能性問題 SAT

- ① 有限集合 $U = \{u_1, u_2, \dots, u_n\}$ に対して、

$$t : U \rightarrow \{T, F\}$$

- ② $t(u_i) = F \iff t(\bar{u}_i) = T$ なる $\bar{U} = \{\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n\}$ に対して、

$$t : U \cup \bar{U} \rightarrow \{T, F\}$$

- ③ $c \subset U \cup \bar{U}$ に対して、 $t(c) = T \iff t(z) = T$ for $\exists z \in c$

- ④ U 上の節集合 C に対して、 $t(C) = T \iff t(c) = T$ for $\forall c \in C$

リテラル $U \cup \bar{U}$ の要素

節 $U \cup \bar{U}$ の部分集合 c ($|c|$: 節の大きさ、要素数)

節集合 節を要素とする集合

充足可能性問題 SAT (続)

事例： 有限集合 U , U 上の節集合 C

質問： $t(C) = T$ となる $t: U \rightarrow \{T, F\}$ が存在するか。

- ① $U = \{u_1, u_2\}$, $C = \{\{u_1, \bar{u}_2\}, \{\bar{u}_1, u_2\}\}$ 「存在する」
 $t(u_1) = t(u_2) = T$ なる t に対して $t(C) = T$ 。
- ② $U = \{u_1, u_2\}$, $C = \{\{u_1, u_2\}, \{u_1, \bar{u}_2\}, \{\bar{u}_1\}\}$ 「存在しない」
 $t(C) = T$ なる t が存在しない。

符号化

決定問題:

入力: 事例 $I \in$ 問題 Π

出力: T または F

符号化 $\varphi : \Pi \rightarrow \{0, 1\}^*$

事例 $I \in \Pi$ を 2 進列に対応

問題 Π の NP 完全性

$\varphi(\Pi)$ についての NP 完全性

符号化の例

$$\Psi := \{0, 1, -, [,], (,), ,\}$$

- ① $k \in \mathbb{Z}$: 非負の数は2進、負の数は前に-。
- ② 1. の値が x : $[x]$ で数値ではなく、識別のための記号。
- ③ (x_1, \dots, x_m) : x_1, \dots, x_m からなる列。

Ψ の要素	0	1	-	[]	()	,
2進列	000	001	010	011	100	101	110	111

- ① $U = \{u_1, u_2\}$, $C = \{\{u_1, \bar{u}_2\}, \{\bar{u}_1, u_2\}\}$:
 $(([1], [10]), (([1], [-10]), ([-1], [10])))$
 を2進列に直した $41 \times 3 = 123$ ビットが入力の長さ
- ② $U = \{u_1, u_2\}$, $C = \{\{u_1, u_2\}, \{u_1, \bar{u}_2\}, \{\bar{u}_1\}\}$:
 $([1], [10]), (([1], [10]), ([1], [-10]), ([-1]))$
 を2進列に直した $48 \times 3 = 144$ ビットが入力の長さ

Cook の定理

Cook の定理

SAT は、 \mathcal{NP} 完全問題である。

(証明略)

Π, Π' : 問題

問題 Π が問題 Π' に多項式時間で還元 $\Pi \propto \Pi'$

言語の多項式時間還元と同様に定義

- ① $\Pi \propto \Pi', \Pi' \propto \Pi'' \implies \Pi \propto \Pi''$
- ② $\Pi \propto \Pi', \Pi' \in \mathcal{P} \implies \Pi \in \mathcal{P}$
- ③ Π が \mathcal{NP} 完全, $\Pi \propto \Pi' \implies \Pi'$ が \mathcal{NP} 完全

3-充足可能性問題 3SAT

事例： 有限集合 U , U 上の大きさ 3 の節からなる節集合 C
 質問： $t(C) = T$ となる $t: U \rightarrow \{T, F\}$ が存在するか。

命題 4.3

3SAT は、NP 完全である。

$U = \{u_1, \dots, u_n\}$, $C = \{c_1, \dots, c_m\}$: SAT の任意の事例
 C と C' の充足可能性が一致するように 3SAT の事例
 $U' := U \cup (\cup_{j=1}^m U'_j)$, $C' := \cup_{j=1}^m C'_j$, $c_j = \{z_1, \dots, z_k\}$

k	U'_j	C'_j
1	$\{y_j^1, y_j^2\}$	$\{\{z_1, y_j^1, y_j^2\}, \{z_1, y_j^1, \bar{y}_j^2\}, \{z_1, \bar{y}_j^1, y_j^2\}, \{z_1, \bar{y}_j^1, \bar{y}_j^2\}\}$
2	$\{y_j^1\}$	$\{\{z_1, z_2, y_j^1\}, \{z_1, z_2, \bar{y}_j^1\}\}$
3	ϕ	$\{c_j\}$
≥ 4	$\{y_j^i \mid 1 \leq i \leq k-3\}$	$\{\{z_1, z_2, y_j^1\}, \{\bar{y}_j^1, z_3, y_j^2\}, \{\bar{y}_j^2, z_4, y_j^3\}, \dots, \{\bar{y}_j^{k-5}, z_{k-3}, y_j^{k-4}\}, \{\bar{y}_j^{k-4}, z_{k-2}, y_j^{k-3}\}, \{\bar{y}_j^{k-3}, z_{k-1}, z_k\}\}$

3-充足可能性問題 3SAT

 $k \geq 4$:

$$t(z_1) = F, \dots, t(z_{l-1}) = F, t(z_l) = T$$

l	i	$t'(y_j^i)$
$l = 1, 2$	$1 \leq i \leq k - 3$	F
$3 \leq l \leq k - 2$	$1 \leq i \leq l - 2$	F
	$l - 1 \leq i \leq k - 3$	T
$l = k - 1, k$	$1 \leq i \leq k - 3$	T

逆に、 $t' : U' \rightarrow \{T, F\}$ が C' を充足させれば、 t' の定義域 U' を U に制限して、 C を充足させる $t : U \rightarrow \{T, F\}$ が得られる。

SAT の事例 U, C が mn の多項式の長さで記述でき、決定的 TM を用いて、その長さの多項式の時間で 3SAT の事例が計算

NP 困難

NP 完全 決定問題に対して定義

NP 困難 決定問題ではなくともよいが、その問題が決定性 TM で多項式時間で解かれると、 $P = NP$ を意味する問題

例:

事例： 有限集合 U , U 上の節集合 C

質問： $t(C) = T$ となる $t: U \rightarrow \{T, F\}$ が存在するとき、その 1 個を出力せよ

確率的推論と最大事後確率設定 \mathcal{NP} の困難性

PR: Probabilistic Reasoning (確率的推論)

MAP: Maximum A Posteriori (最大事後確率)

確率的推論 (PR)

事例: $P_{X_1 \dots X_N}(x_1, \dots, x_N), x_1 \in \mathcal{X}_1, \dots, x_N \in \mathcal{X}_N, i \in \mathcal{N}$

質問: $P_{X_i}(x_i), x_i \in \mathcal{X}_i$

最大事後確率設定 (MAP)

事例: $P_{X_1 \dots X_N}(x_1, \dots, x_N), x_1 \in \mathcal{X}_1, \dots, x_N \in \mathcal{X}_N$

質問: $\operatorname{argmax}_{x_1, \dots, x_N} P_{X_1 \dots X_N}(x_1, \dots, x_N)$

確率的推論と最大事後確率設定 \mathcal{NP} の困難性

確率的推論の決定問題版 (PRD)

事例： $P_{X_1 \dots X_N}(x_1, \dots, x_N)$, $x_1 \in \mathcal{X}_1, \dots, x_N \in \mathcal{X}_N, i \in \mathcal{N}$,
 $x_i \in \mathcal{X}_i$

質問： $P_{X_i}(x_i) > 0$ が成立するか

最大事後確率設定の決定問題版 (MAPD)

事例： $P_{X_1 \dots X_N}(x_1, \dots, x_N)$, $x_1 \in \mathcal{X}_1, \dots, x_N \in \mathcal{X}_N$, 有理数
 $0 < p < 1$

質問： $P_{X_1 \dots X_N}(x_1, \dots, x_N) > p$ なる
 $x_1 \in \mathcal{X}_1, \dots, x_N \in \mathcal{X}_N$ は存在するか。

PR は \mathcal{NP} 困難

定理 4.3

PRD は \mathcal{NP} 完全、すなわち PR は \mathcal{NP} 困難

証明: 3SAT からの多項式時間還元性のみを示す。

$U = \{u_1, \dots, u_n\}$, $C = \{c_1, \dots, c_m\}$: 3SAT の事例

C が U 上充足可能 \iff ある $i \in \mathcal{N}$ と $x_i \in \mathcal{X}_i$ について $P_{X_i}(x_i) > 0$

なる $P_{X_1 \dots X_N}(x_1, \dots, x_N)$ を構成

PR は NP 困難 (続)

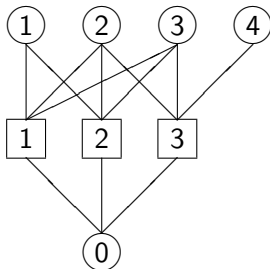
$\mathcal{N} = \{0, 1, \dots, n\}$ (変数頂点集合)

$\mathcal{M} := \{1, \dots, m\}$ (因数頂点集合)

$\mathcal{N}(j)$: c_j , $j \in \mathcal{M}$ のリテラル u_i の 3 個の添え字 $i \in \mathcal{N}$ と 0 からなる集合
 $j \in \mathcal{M}$ と $i \in \mathcal{N}(j)$ を辺として結んで因数グラフを構成

$\mathcal{M}(i)$, $i \in \mathcal{N}$ が定義

例 4.10 $m = 3$, $n = 4$



PR は NP 困難 (続)

$$f_j(u_0, c_j) = \begin{cases} 1, & u_0 = \text{F, or, } c_j = \text{T} \\ 0, & \text{otherwise} \end{cases}$$

$$P_{U_0 \dots U_n}(u_0, \dots, u_n) = K \prod_{j \in \mathcal{M}} f_j(u_0, c_j) \quad (21)$$

$$K := \left[\sum_{u_0, \dots, u_n} \prod_{j \in \mathcal{M}} f_j(u_0, c_j) \right]^{-1}$$

PR は NP 困難 (続)

C が U 上充足可能

$$\Rightarrow c_1 = \dots = c_m = T$$

なる $u_1, \dots, u_n \in \{T, F\}$ が存在

$$\Rightarrow \text{各 } j \in \mathcal{M} \text{ について, } f_j(u_0 = T, c_j) = 1$$

なる $u_i \in \{T, F\}, i \in \mathcal{N}(j) \setminus \{0\}$ が存在

$$\Rightarrow P_{U_0 \dots U_n}(u_0 = T, u_1, \dots, u_n) \propto \prod_{j \in \mathcal{M}} f_j(u_0 = T, c_j) = 1$$

なる $u_i \in \{T, F\}, i \in \mathcal{N} \setminus \{0\}$ が存在

$$\Rightarrow P_{U_0}(u_0 = T) > 0$$

PR は NP 困難 (続)

C が U 上充足可能でない

$\Rightarrow c_1 = \dots = c_m = T$ となる $u_1, \dots, u_n \in \{T, F\}$ が存在しない

\Rightarrow 任意の $u_i \in \{T, F\}$, $i \in \mathcal{N}(j) \setminus \{0\}$ について

$f_j(u_0 = T, c_j) = 0, j \in \mathcal{M}$ が存在

\Rightarrow 任意の $u_i \in \{T, F\}$, $i \in \mathcal{N} \setminus \{0\}$ について

$P_{U_0 \dots U_n}(u_0 = T, u_1, \dots, u_n) \propto \prod_{j \in \mathcal{M}} f_j(u_0 = T, c_j) = 0$

$\Rightarrow P_{U_0}(u_0 = T) = 0$

3SAT の事例 U, C が mn の多項式の長さで記述でき、決定的 TM を用いて、その長さの多項式の時間で PRD の事例が計算

例 4.10

$U = \{u_1, u_2, u_3, u_4\}$, $C = \{\{u_1, u_2, u_3\}, \{\bar{u}_1, \bar{u}_2, u_3\}, \{u_2, \bar{u}_3, u_4\}\}$ であれば, $\mathcal{N} = \{1, 2, 3, 4\}$, $\mathcal{M} = \{1, 2, 3\}$ $\mathcal{N}(1) = \{0, 1, 2, 3\}$, $\mathcal{N}(2) = \{0, 1, 2, 3\}$, $\mathcal{N}(3) = \{0, 2, 3, 4\}$, $u_0 \neq *$ のとき,

$$f_1(u_0, u_1, u_2, u_3) = \begin{cases} 1, & u_0 = F \text{ または } c_1 = T, \\ 0, & \text{それ以外,} \end{cases}$$

$$f_2(u_0, u_1, u_2, u_3) = \begin{cases} 1, & u_0 = F \text{ または } c_2 = T, \\ 0, & \text{それ以外,} \end{cases}$$

$$f_3(u_0, u_2, u_3, u_4) = \begin{cases} 1, & u_0 = F \text{ または } c_3 = T \\ 0, & \text{それ以外} \end{cases}$$

$$P_{U_0 U_1 U_2 U_3 U_4}(u_0, u_1, u_2, u_3, u_4) \\ \propto f_1(u_0, u_1, u_2, u_3) f_2(u_0, u_1, u_2, u_3) f_3(u_0, u_2, u_3, u_4)$$

例 4.10

$$C(u_1, u_2, u_3, u_4) = \text{T} \iff f_1(u_0 = \text{T}, u_1, u_2, u_3) f_2(u_0 = \text{T}, u_1, u_2, u_3) \cdot f_3(u_0 = \text{T}, u_2, u_3, u_4) = 1$$

u_0	u_1	u_2	u_3	u_4	c_1	c_2	c_3	C	f_1	f_2	f_3	$f_1 f_2 f_3$
F	F	F	F	F	F	T	T	F	1	1	1	1
F	F	F	F	T	F	T	T	F	1	1	1	1
F	F	F	T	F	T	T	F	F	1	1	1	1
F	F	F	T	T	T	T	T	T	1	1	1	1
F	F	T	F	F	T	T	T	T	1	1	1	1
F	F	T	F	T	T	T	T	T	1	1	1	1
F	F	T	T	F	T	T	T	T	1	1	1	1
F	F	T	T	T	T	T	T	T	1	1	1	1
F	T	F	F	F	T	T	T	T	1	1	1	1
F	T	F	F	T	T	T	T	T	1	1	1	1
F	T	F	T	F	T	T	F	F	1	1	1	1
F	T	F	T	T	T	T	T	T	1	1	1	1
F	T	T	F	F	T	F	T	F	1	1	1	1
F	T	T	F	T	T	F	T	F	1	1	1	1
F	T	T	T	F	T	T	T	T	1	1	1	1
F	F	T	T	T	T	T	T	T	1	1	1	1
T	F	F	F	F	F	T	T	F	0	1	1	0
T	F	F	F	T	F	T	T	F	0	1	1	0
T	F	F	T	F	T	T	F	F	1	1	0	0

PR が解ければ、SAT が解ける

アルゴリズム 4.1 を適用:

- ① $n_{i \rightarrow j}(u_i) := 1, m_{j \rightarrow i}(u_i) := 1, i \in \mathcal{N}, j \in \mathcal{M}$
- ② $j \in \mathcal{M}$ について,

$$m_{j \rightarrow i}(u_i) := \sum_{u_l, l \in \mathcal{N}(j) \setminus \{i\}} f_j(u_0, c_j) \prod_{k \in \mathcal{N}(j) \setminus \{i\}} n_{k \rightarrow j}(u_k),$$

$$n_{i \rightarrow j}(u_i) := \prod_{k \in \mathcal{M}(i) \setminus \{j\}} m_{k \rightarrow i}(u_i).$$

- ③ $b_0(u_0) \propto \prod_{j \in \mathcal{M}} m_{j \rightarrow 0}(u_0)$

2., 3. を繰り返す。

系 4.1

系 4.1, Shimony, 1994

MAPD は \mathcal{NP} 完全 , すなわち MAP は \mathcal{NP} 困難である。証明: 定理 4.3 の証明で , (21) の K の値は高々 2^{n+1} なので ,

$$P_{U_0 \dots U_n}(u_0 = \text{T}, u_1, \dots, u_n) \begin{cases} \geq 2^{-n-1}, & u_1, \dots, u_m \text{ satisfiable} \\ = 0, & u_1, \dots, u_m \text{ non-satisfiable} \end{cases}$$

$$p := 2^{-n-1}$$

\mathcal{NP} 困難の注意点

- \mathcal{NP} 困難性は、すべての場合に計算時間がかかることを意味しない
- 問題の中に計算時間がかかる事例が含まれていることのみを主張
- 条件を加えれば、 \mathcal{NP} 困難性が崩れることがある

例: PRD や MAPD でも, 因子グラフが巡回経路を含まない場合, \mathcal{P} 問題

接合木における \mathcal{NP} 困難問題接合木における \mathcal{NP} 困難問題 (第2章)

$\max_{C \in \mathbf{E}_T} |C|$ を最小にする $T = (\mathbf{U}_T, \mathbf{E}_T)$ を見いだす問題が \mathcal{NP} 困難

Partial K -Tree が \mathcal{NP} 完全であることによる

事例 : $G = (\mathbf{U}, \mathbf{E})$ および正整数 K

質問 : 以下を満足する chordal な $G' = (\mathbf{U}, \mathbf{E}')$ が存在するか。

- (1) $\mathbf{E} \subseteq \mathbf{E}'$
- (2) G' の各クリークの頂点数が高々 K 以下。

- 無向グラフに辺を加えて chordal にすることはいつでも可能
- 任意の無向グラフで、各頂点の最大値を最小にする接合木がわかれば、Partial K -Tree の任意の K について解けたことになる。