

避難行動における個人間異質性を捉える為のパラメータ推定法 Parameter estimation method to capture heterogeneity in evacuation behavior

荒木 雅弘¹⁾, Giancarlo Troncoso Parady²⁾
Masahiro ARAKI, Giancarlo Troncoso PARADY

1)東京大学大学院 工学系研究科, 博士後期課程2年, 修士(工学), The University of Tokyo, Graduate Student, Mr. (Eng.)

2)東京大学大学院 工学系研究科, 講師, 博士(工学), The University of Tokyo, Lecturer, Dr. (Eng.)

避難行動, 個人間異質性, 潜在クラスモデル, DAEM algorithm
Evacuation Behavior, Heterogeneity, Latent Class Model, Deterministic Annealing EM algorithm

1. はじめに

防災・減災の施策の理論的根拠の獲得の為に, 避難行動に関する調査[1]を行い, データを用いた定量的分析による知見の発見を行うことの重要性は論を俟たない. 防災・減災の施策は人命を直接的に預かる以上, 施策の理論的根拠となるパラメータは十分に信頼に足る必要があり, その信頼性を向上させていくことは, 必要不可欠な研究課題の一つである.

この一端として, 本研究では, 潜在クラスモデル (LCM/Latent Class Model) [2]のパラメータの推定性能の向上に取り組んだ. 避難行動においては個人間異質性が大きいと想像されるが, 是を表現する古典的モデルの一つとして LCM があり, EM algorithm[3]を用いた推定が一般的に行われている. しかし, LCM の尤度関数は非凸で多峰性を有し, EM algorithm のように初期値から単調に峰を登る最適化手法では, 局所解に陥る可能性が高い. この為, 多くの初期値から試行する多出発局所探索法が一般的に用いられているが, クラス数を増やす場合のように, 多峰性が増す場合の解の収束性については, 議論が乏しい.

本研究では, 平衡統計力学[4]の概念を援用して, Deterministic Annealing EM (DAEM) algorithm [5]を整合的に再導出した上で, 避難開始選択行動に関する実データを用いた分析を行い, 解の収束性に関する EM algorithm との比較検証と, 避難行動に関する信頼性の高い知見の発見を行った.

2. 推定法

DAEM algorithm の導出を以下に簡単に示す.

まず, 非観測変数が離散値であることと, データ集合が i.i.d. 条件を満たすことを仮定する. なお, LCM はこの仮定を満たす. この時, 表 1 のように, LCM における i.i.d. なデータ集合の一つ n (ある意思決定者 n) を, 平衡統計力学における相互独立な部分系 n (図 1) と見做すことが出来る. 部分系 n について, 熱力学的物理量のミクロ表記を与えると, 表 2 のようになる. 等温条件下での平衡状態, すなわち自由エネルギーが最小になること (変分原理) を考えると, 確率分布は, 制約条件 $\sum_{c \in C} p_c = 1$ のもとでラグランジュ未定乗数法を解いて, 式(1)のようになる. なお, この場合の確率分布 p_c はカノニカル分布, 自由エネルギーはヘルムホルツの自由エネルギーに他ならず, 熱力学的原理により, このときエントロピー最大化も成立している.

$$p_c = \frac{\exp(-\beta E_c)}{\sum_{c' \in C} \exp(-\beta E_{c'})}, \text{ where } \beta := \frac{1}{k_B T} \quad \text{式(1)}$$

次に, 相互独立な部分系 n の集合としての全系 N に関して, 熱力学的物理量は部分系のその総和となるので, 式(2)~(4)が成立する. また, 式変形を続けていくと, 式(5)のように, ヘルムホルツの自由エネルギーの分配関数表記が成立することが確認できる.

$$F(\beta) = \sum_{n \in N} F_n(\beta) = U - \frac{1}{\beta} S \quad \text{式(2)}$$

$$U = \sum_{n \in N} U_n \quad \text{式(3)}$$

$$S = \sum_{n \in N} S_n \quad \text{式(4)}$$

$$F(\theta, \beta) = -\frac{1}{\beta} \ln \left\{ \sum_{x_{mis}} p(X_{obs}, X_{mis} | \theta)^\beta \right\} \quad \text{式(5)}$$

そして, 式(5)より, ヘルムホルツの自由エネルギーの最小化問題が, 不完全データの対数尤度関数の最大化問題と, $\beta = 1$ で逆符号をとれば等価になることが明らかとなった.

等温条件下の平衡状態では, ヘルムホルツの自由エネルギーを小さくするように, 「内部エネルギーを下げる傾向」と「エントロピーを増大する傾向」が拮抗しており, DAEM algorithm (図 2, 図 3) では, 温度を徐々に下げつつ, この二つを交互に行うことで, パラメータを求めていると見做すことが出来る. なお, EM algorithm はこの特殊形 ($\beta = 1$ のみの場合) と見做せる.

3. モデル及びデータ, 推定結果

本研究では, 避難開始選択行動を分析する為に, 各クラスの離散選択モデルとして, Sequential Logit Model[6]を使用した.

データとして, 2011 年 3 月 11 日の東北地方太平洋沖地震発災当日の避難行動データ[1]及び, 各種一次資料及び各行政・研究機関への照会によって収集・構築した状況データを用いた.

推定結果の解釈の一例を述べる. 図 4 より, EM algorithm と比較して一意に決まる DAEM algorithm が必ずしも良い最終尤度を返すわけではないが, $c = 4$ のようにクラス数が増えるほど局所解の数が増大するなかで, 良い最終尤度の解を探索可能なことが分かる. 表 3 より, 避難行動において個人間異質性は確かに存在し, 例えば避難指示が避難を促進するクラス (class 1, 2) と, 避難を促進しないクラス (class 3, 4) が確認された.

紙幅の都合上, 推定結果及び検証結果, そして避難行動に関する解釈と知見については, 当日に詳細に発表する予定である.

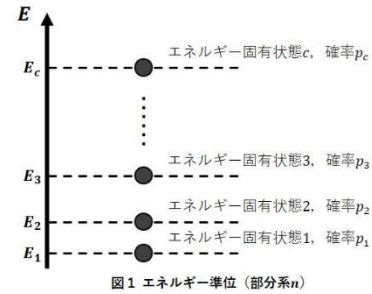
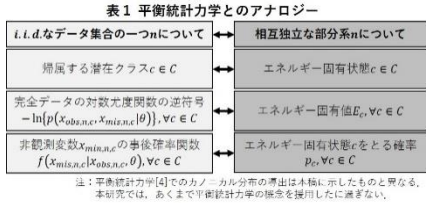


図 1 エネルギー単位 (部分系n)

表 2 部分系nの熱力学的物理量

部分系nの熱力学的物理量	同左 (ミクロ表記)
シャノンエントロピー S_n	$S_n := -k_B \sum_{c \in C} p_c \ln p_c$
内部エネルギー U_n	$U_n := \sum_{c \in C} p_c E_c$
温度Tでの自由エネルギー $F_n(T)$	$F_n(T) := U_n - TS_n$

注： k_B はボルツマン定数という定数であり、導出に影響しない。

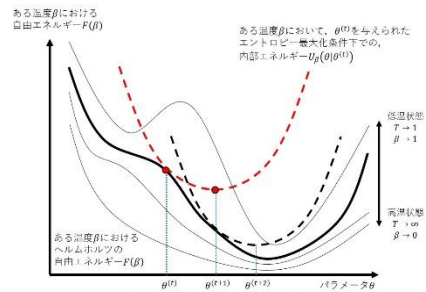


図 2 DAEM algorithmの概念図

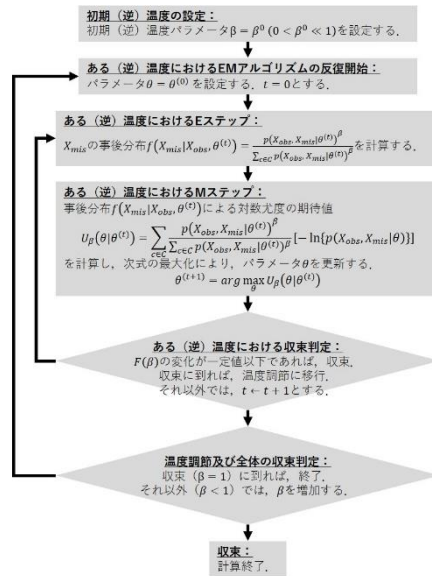


図 3 DAEM algorithmのフロー

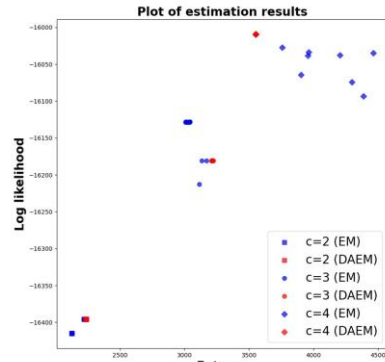


図 4 推定結果

注：選択確率のパラメータ β 及び、クラス固有確率のパラメータ p_i は、 $[-1, +1]$ の範囲からランダムに初期値を与えた。その初期条件で各20回試行した結果を図に示している。ただし、厳密研究に資及がある通り、大きいクラス数 (i.e. $c = 4$) においてはEM algorithm (not DAEM algorithm) が収束しない可能性が確認され、これに関する完全な真実の特定は完了していない。特定を完了し、当日にはまとめて発表する予定である。

参考文献

- [1] 国土交通省, 復興支援調査アーカイブ

表 3 推定結果

Estimation Result of c=4 (DAEM)	Class 1	Class 2	Class 3	Class 4
説明変数	推定値 (z値)	推定値 (z値)	推定値 (z値)	推定値 (z値)
避難選択確率 (+: 避難, -: 非避難)				
本県の計測地震動	1.114 (6.665***)	-0.261 (-3.783***)	0.671 (5.739***)	-8.599 (-9.036***)
地震回数: 震度 1, 2	0.093 (11.663***)	0.042 (8.907***)	0.726 (18.012***)	-2.143 (-11.980***)
地震回数: 震度 3	-0.207 (-7.859***)	0.181 (18.471***)	0.205 (5.589***)	5.041 (12.029***)
地震回数: 震度 4	-0.063 (-1.548)	0.048 (2.518**)	0.153 (1.659*)	10.653 (11.918***)
地震回数: 震度 5	-0.276 (-1.124)	-1.306 (-11.516***)	-13.956 (-0.069)	-1.228 (-2.215**)
大津波警報の津波予報高さ (1/10)	0.461 (3.696***)	0.119 (2.033**)	0.455 (2.652**)	7.249 (9.940***)
避難指示ダミー	0.327 (2.953***)	0.199 (4.423***)	-0.353 (-3.899***)	-4.475 (-8.838***)
物理被害ダミー	-0.129 (-1.085)	0.168 (3.547***)	0.009 (0.096)	4.840 (8.708***)
所在 (非避難) に関する定数項	10.286 (11.367***)	1.572 (4.302***)	4.327 (6.925***)	-15.973 (-4.423***)
クラス固有確率				
海沿からの距離 (1/1000)	0.0 (---)	-0.518 (-16.279***)	-1.650 (-21.974***)	-0.764 (-11.575***)
標高 (1/10)	0.0 (---)	-0.936 (-19.306***)	-1.891 (-21.101***)	-1.523 (-11.029***)
(性別) 女性ダミー	0.0 (---)	0.222 (3.028***)	0.653 (6.850***)	0.124 (1.064)
(年齢) 30歳代ダミー	0.0 (---)	-0.258 (-1.611)	-0.625 (-2.777***)	-0.907 (-3.304***)
(年齢) 40歳代ダミー	0.0 (---)	-0.313 (-1.954*)	-1.028 (-4.529***)	-0.721 (-2.758***)
(年齢) 50歳代ダミー	0.0 (---)	-0.251 (-1.615)	-0.602 (-2.812***)	-0.936 (-3.672***)
(年齢) 60歳代ダミー	0.0 (---)	-0.152 (-0.979)	-0.343 (-1.630)	-0.383 (-1.600)
(年齢) 70歳代ダミー	0.0 (---)	-0.066 (-0.405)	0.125 (0.583)	-0.271 (-1.104)
(職業) オフィスワーカーダミー	0.0 (---)	-0.641 (-6.584***)	-0.714 (-4.952***)	-1.167 (-5.890***)
(職業) 家庭ダミー	0.0 (---)	0.388 (4.444***)	0.794 (6.997***)	0.383 (2.806**)
(職業) 水産業従事者ダミー	0.0 (---)	-0.447 (-2.209**)	0.624 (2.802**)	0.855 (3.428***)
(同居家族) 乳幼児ダミー	0.0 (---)	-0.118 (-1.389)	-0.468 (-3.999***)	-0.142 (-0.995)
(同居家族) 高齢者・身体障害者ダミー	0.0 (---)	-0.088 (-1.236)	-0.206 (-2.267**)	-0.261 (-2.299**)
(防災意識) 三階階層ダミー	0.0 (---)	1.869 (19.343***)	3.194 (26.672***)	0.951 (6.200**)
(防災意識) ハザードマップの知識ダミー	0.0 (---)	-0.106 (-1.352)	-0.024 (-0.242)	-0.358 (-2.796***)
(防災意識) 有教等の避難経験ダミー	0.0 (---)	-0.057 (-0.693)	0.001 (0.013)	0.222 (1.703*)
(防災意識) 震災前の災害対策ダミー	0.0 (---)	0.268 (3.572**)	0.359 (3.589***)	0.086 (0.709)
(防災意識) 避難所の知識ダミー	0.0 (---)	0.248 (3.122**)	0.805 (6.981***)	0.721 (4.947***)
定数項 (関値)	0.0 (---)	1.358 (7.820***)	0.118 (0.488)	0.429 (1.527)
クラスサイズ	0.238	0.479	0.215	0.067
サンプル数	7418			
初期対数尤度 (無情報モデル)	-20567.063			* : 10%有意 (≧1.645)
最終対数尤度	-16009.423			** : 5%有意 (≧1.960)
McFaddenの決定係数	0.222			***: 1%有意 (≧2.576)
修正済み決定係数	0.217			
赤池情報量規準 (AIC)	32204.847			
ベイズ情報量規準 (BIC)	32847.632			

- [2] Bhat, C. R. (1997). An endogenous segmentation mode choice model with an application to intercity travel. Transportation science, 31(1), 34-48.

- [3] Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological), 39(1), 1-22.

- [4] 田崎晴明. (2008). 統計力学 (No. 37-38). 培風館.

- [5] Ueda, N., & Nakano, R. (1998). Deterministic annealing EM algorithm. Neural networks, 11(2), 271-282.

- [6] Fu, H., & Wilmot, C. G. (2004). Sequential logit dynamic travel demand model for hurricane evacuation. Transportation Research Record, 1882(1), 19-26.