集中理論談話会 #9

Bhat, C.R., Sidharthan, R.: A simulation evaluation of the maximum approximate composite marginal likelihood (MACML) estimator for mixed multinomial probit models, *Transportation Research Part B: Methodological*, Vol.45, No.7, pp.940-953, 2011.

2014/06/28

柳沼秀樹

1. MNP概要(MNLとの比較を通じて)

多項ロジットモデル

$$P(i)=\frac{\exp(\mu V_i)}{\sum_{j\in C}\exp(\mu V_j)}$$

- Closed-form
- 選択肢間の相関を無視(IIA)

表現力は劣るが、通常の最適化アルゴリズムでパラメータ推定が可能であり、ソースコードの開発が容易

⇒爆発的に各分野に普及

多項プロビットモデル

$$P(i)=\int_{\varepsilon_1=-\infty}^{\varepsilon_i+V_i-\varepsilon_1}\cdots\int_{\varepsilon_i=-\infty}^{\infty}\cdots\int_{\varepsilon_J=-\infty}^{\varepsilon_i+V_i-\varepsilon_J}\phi(\varepsilon)d\varepsilon_J\cdots d\varepsilon_1$$

$$\phi(\varepsilon)=\frac{1}{\left(\sqrt{2\pi}\right)^{J-1}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}\varepsilon\Sigma^{-1}\varepsilon'\right)$$

- Open-form
- 選択肢間の相関を表現

表現力は高いが、選択肢数−1の多重積分が必要であり、パラメータ推定は非常に煩雑

⇒計算が困難であるため
近年まで伸び悩む…

2. MNPの推定方法

開発当初から多重積分の計算がネックとなり、各種の数値計算アルゴリズムが提案

	選択肢数	手法	
1970			
↓	3～4	• 級数近似法 • Clarkの近似法etc…	解析的近似
1990			
↓	3～10	• シミュレーション法 • ベイズ推定	乱数近似
	10以上		
2010		• MACML	
↓			

1. 論文の概要

(1) Maximum Approximate Composite Marginal Likelihood (MACML) estimationの提案

- ✓ Open-fromな離散選択モデル(c.f. MNP,MXL)のパラメータを 簡単かつ高速で推定する手法を構築
- ✓ MACML推定は2つのテクニックにより構成
  - ①多変量累積標準正規分布（MVNCD※1）の解析的近似手法
  - ②合成周辺尤度(CML※2)を用いたパラメータ推定

(2) 各種（Mixed）Probitモデルへの適用方法の提示

- ✓ Cross-section, Panel, Spatial Correlation etc…

(3) 数値実験による有効性の確認

- ✓ 通常の推定手法(MSL)と比較して，計算時間は約38倍速く(66.09→1.96)，推定値のバイアスは7.3ポイント低い（9.8%→2.5%）.

※1 MVNCD: Multi-Variate standard Normal Cumulative Distribution
※2CML: Composite Marginal Likelihood

1. MACML推定に関する論文

① Bhat, C.R.: The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models, *Transportation Research Part B: Methodological*, Vol.45, No.7, pp.923-939, 2011.

⇒MACMLの提案とProbitへの適用方法

② Bhat, C.R., Sidharthan, R.: A simulation evaluation of the maximum approximate composite marginal likelihood (MACML) estimator for mixed multinomial probit models, *Transportation Research Part B: Methodological*, Vol.45, No.7, pp.940-953, 2011.

⇒数値実験によるMACMLの性能評価

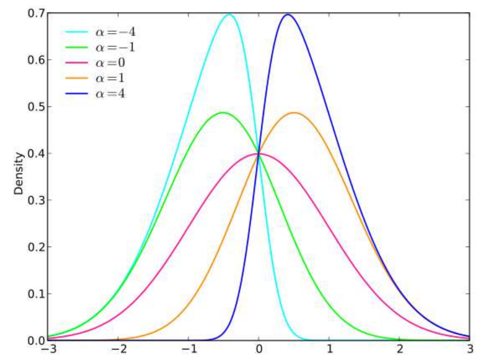
本日の内容

① Bhat, C.R., Sidharthan, R.: A new approach to specify and estimat normally mixed multinomial probit models, *Transportation Research Part B: Methodological*, Vol.46, No.7, pp.817-833, 2012.

⇒Skew-Normal(歪度≠0)なMNPへの適用

② Bhat, C.R., Dubey, S.: A new estimation approach to integ psychological constructs in choice modeling, *Transportati Methodological*, Vol.67, pp.68-85, 2014.

⇒潜在クラスモデルへの適用



2. MACMLアプローチ(1)

2.1 MVNCD(多変量累積標準正規分布)の解析的近似

⇒多変量正規分布を単変量分布の積で近似

【設定1: 分布の分解】

$$\Pr(\boldsymbol{W} < \boldsymbol{w}) = \Pr(W_1 < w_1, W_2 < w_2, W_3 < w_3, \ldots, W_I < w_I).$$

W: 多変量標準正規分布

同時確率を下記のように分布の積に分解

$$\Pr(\boldsymbol{W} < \boldsymbol{w}) = \underbrace{\Pr(W_1 < w_1, W_2 < w_2)}_{\text{二変量周辺分布}} \times \prod_{i=3}^{I} \underbrace{\Pr(W_i < w_i | W_1 < w_1, W_2 < w_2, W_3 < w_3, \ldots, W_{i-1} < w_{i-1})}_{\text{単変量条件つき分布(I>3)}}.$$

【設定2:インディケータIでの分散共分散表現 】

$$\tilde{\mathbf{I}}_i = \begin{cases} 1 & W_i < w_i \\ 0 & otherwise \end{cases}$$

$$E(\tilde{I}_i) = \Phi(w_i)$$

Iの期待値を単変量累積標準正規分布**Φ** で評価

$$Cov(\tilde{I}_i, \tilde{I}_i) = Var(\tilde{I}_i) = \Phi(w_i) - \Phi^2(w_i) = \Phi(w_i)[1 - \Phi(w_i)],$$

$$\mathrm{Cov}(\tilde{I}_i, \tilde{I}_j) = E(\tilde{I}_i\tilde{I}_j) - E(\tilde{I}_i)E(\tilde{I}_j) = \Phi_2(w_i, w_j, \rho_{ij}) - \Phi(w_i)\Phi(w_j), i \neq j$$

上記を統合

$$\Pr(W_i < w_i | W_1 < w_1, W_2 < w_2, W_3 < w_3, \ldots, W_{i-1} < w_{i-1}) = E(\tilde{I}_i | \tilde{I}_1 = 1, \tilde{I}_2 = 1, \tilde{I}_3 = 1, \ldots, \tilde{I}_{i-1} = 1).$$

2. MACMLアプローチ(2)

【線形回帰モデルでの展開】

$$\Pr(W_i < w_i | W_1 < w_1, W_2 < w_2, W_3 < w_3, \ldots, W_{i-1} < w_{i-1}) = \underline{E(\tilde{I}_i | \tilde{I}_1 = 1, \tilde{I}_2 = 1, \tilde{I}_3 = 1, \ldots, \tilde{I}_{i-1} = 1)}.$$

$$\underline{\tilde{I}_i - E(\tilde{I}_i) = \boxed{\boldsymbol{\alpha}'}[\tilde{\boldsymbol{I}}_{<i} - E(\tilde{\boldsymbol{I}}_{<i})] + \tilde{\eta},}$$ 誤差項

※ $\tilde{\boldsymbol{I}}_{<i} = (\tilde{I}_1, \tilde{I}_2, \ldots \tilde{I}_{i-1})$

$$\boxed{\hat{\boldsymbol{\alpha}}} = \boldsymbol{\Omega}_{<i}^{-1} \cdot \boldsymbol{\Omega}_{i,<i},$$ 最小二乗係数

$$\boldsymbol{\Omega}_{<i} = \mathrm{Cov}(\boldsymbol{I}_{<i}, \boldsymbol{I}_{<i}) = \begin{bmatrix} \mathrm{Cov}(\tilde{I}_1, \tilde{I}_1) & \mathrm{Cov}(\tilde{I}_1, \tilde{I}_2) & \mathrm{Cov}(\tilde{I}_1, \tilde{I}_3) & \cdots & \mathrm{Cov}(\tilde{I}_1, \tilde{I}_{i-1}) \\ \mathrm{Cov}(\tilde{I}_2, \tilde{I}_1) & \mathrm{Cov}(\tilde{I}_2, \tilde{I}_2) & \mathrm{Cov}(\tilde{I}_2, \tilde{I}_3) & \cdots & \mathrm{Cov}(\tilde{I}_2, \tilde{I}_{i-1}) \\ \mathrm{Cov}(\tilde{I}_3, \tilde{I}_1) & \mathrm{Cov}(I_3, I_2) & \mathrm{Cov}(\tilde{I}_3, \tilde{I}_3) & \cdots & \mathrm{Cov}(\tilde{I}_3, \tilde{I}_{i-1}) \\ \vdots & & & & \\ \mathrm{Cov}(\tilde{I}_{i-1}, \tilde{I}_1) & \mathrm{Cov}(\tilde{I}_{i-1}, \tilde{I}_2) & \mathrm{Cov}(\tilde{I}_{i-1}, \tilde{I}_3) & \cdots & \mathrm{Cov}(\tilde{I}_{i-1}, \tilde{I}_{i-1}) \end{bmatrix}, \quad \boldsymbol{\Omega}_{i,<i} = \mathrm{Cov}(\boldsymbol{I}_{<i}, \boldsymbol{I}_i) = \begin{bmatrix} \mathrm{Cov}(\tilde{I}_i, \tilde{I}_1) \\ \mathrm{Cov}(\tilde{I}_i, \tilde{I}_2) \\ \mathrm{Cov}(\tilde{I}_i, \tilde{I}_3) \\ \vdots \\ \mathrm{Cov}(\tilde{I}_i, \tilde{I}_{i-1}) \end{bmatrix}.$$

【単変量正規分布での近似】

$$\Pr(W_i < w_i | W_1 < w_1, W_2 < w_2, \ldots, W_{i-1} < w_{i-1}) \boxed{\approx \Phi(w_i) + (\boldsymbol{\Omega}_{<i}^{-1} \cdot \boldsymbol{\Omega}_{i,<i})'(1 - \Phi(w_1), 1 - \Phi(w_2) \ldots 1 - \Phi(w_{i-1}))'}$$

多変量正規分布を選択肢数-1の単変量正規分布で表現
⇒計算量は大幅に削減される！

2. MACMLアプローチ(2)

2.2 CML（合成周辺尤度）を用いたパラメータ推定

⇒尤度関数を選択結果別の部分尤度（周辺尤度）の積で表現

【合成周辺尤度　※パネルを想定】

全情報尤度関数（通常）

$$L^1_{CML}(\boldsymbol{\theta}, \boldsymbol{m}) = L(\boldsymbol{\theta}, \boldsymbol{m}) = \mathrm{Prob}(C_1 = m_1, C_2 = m_2, C_3 = m_3, \ldots, C_T = m_T).$$

合成周辺尤度関数

$$L^2_{CML}(\boldsymbol{\theta}, \boldsymbol{m}) = \mathrm{Prob}(C_1 = m_1) \times \mathrm{Prob}(C_2 = m_2) \times \mathrm{Prob}(C_3 = m_3) \times \ldots \times \mathrm{Prob}(C_T = m_t)$$

$$L^3_{CML}(\boldsymbol{\theta}, \boldsymbol{y}) = \prod_{t=1}^{T-1} \prod_{w=t+1}^{T} \mathrm{Prob}(C_t = m_t, C_w = m_w)$$

C_t: t期における個人の選択
m_t: t期における個人の選択結果

【性質】

CML推定量は一致性と漸近正規性を持つ

全情報尤度を周辺尤度に分割する事は、統計的な望ましさを保持しつつ計算量が大幅に削減される

3. 数値実験による性能評価

対象モデル

Cross-section random coefficients model (Mixed MNP)

$$U_{qi} = \boldsymbol{\beta}_{\boldsymbol{q}}' \boldsymbol{x}_{\boldsymbol{qi}} + \varepsilon_{qi} \qquad \boldsymbol{\beta}_{\boldsymbol{q}} \sim MVN(\boldsymbol{b}, \boldsymbol{\Omega})$$

$$L_q = \int_{\beta=-\infty}^{\infty} \left\{ \int_{\lambda=-\infty}^{\infty} \left(\prod_{i \neq m} \left[\Phi \left\{ \left[-\sqrt{2} (\boldsymbol{\beta}' \boldsymbol{z}_{\boldsymbol{qim}}) \right] + \lambda \right\} \right] \right) \phi(\lambda) d\lambda \right\} f(\boldsymbol{\beta} | \boldsymbol{b}, \boldsymbol{\Omega}) \boldsymbol{d\beta},$$

where $\boldsymbol{z}_{\boldsymbol{qim}} = \boldsymbol{x}_{\boldsymbol{qi}} - \boldsymbol{x}_{\boldsymbol{qm}}$

q: 個人
i : 選択肢
ε: 誤差項 IIDガンベル

真値の設定

$\boldsymbol{b}$ = (1.5, −1, 2, 1, −2)

$$\boldsymbol{\Omega} = \begin{bmatrix} 1 & -0.50 & 0.25 & 0.75 & 0 \\ -0.50 & 1 & 0.25 & -0.50 & 0 \\ 0.25 & 0.25 & 1 & 0.33 & 0 \\ 0.75 & -0.50 & 0.33 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5000人分の仮想データを乱数を用いて20セットの実験データを作成

4. 性能比較結果 MSL vs MACML(1)

Cross-sectional random coefficients model

【Diagonal case】

・分散共分散行列の対角成分のみを推定

・計算時間:平均で約34倍速く、かつ速度にバラツキがない(安定的)

・バイアス: 平均で7.3ポイント低く、分散成分についても良好

Table 1a
Evaluation of the ability to recover true parameters for the cross-sectional diagonal case.

Parameter	True value	MSL method					MACML method				
		Parameter estimates		Standard error estimates			Parameter estimates		Standard error estimates		
		Mean estimate	Absolute percentage bias (%)	Asymptotic standard error	Simulation standard error	Simulation adjusted asymptotic standard error	Mean estimate	Absolute percentage bias (%)	Asymptotic standard error	Approximation standard error	Approximation adjusted asymptotic standard error
Mean values of the β vector											
*b*1	1.500	1.366	9.0	0.129	0.050	0.139	1.472	1.9	0.167	0.022	0.169
*b*2	−1.000	−0.906	9.4	0.089	0.033	0.095	−0.976	2.4	0.113	0.014	0.114
*b*3	2.000	1.801	10.0	0.167	0.066	0.180	1.940	3.0	0.218	0.028	0.219
*b*4	1.000	0.906	9.4	0.089	0.034	0.095	0.977	2.3	0.114	0.014	0.114
*b*5	−2.000	−1.820	9.0	0.170	0.067	0.182	−1.960	2.0	0.220	0.028	0.222
Standard deviations of the β vector											
$\sigma 1$	1.000	0.885	11.5	0.111	0.038	0.117	0.958	4.2	0.135	0.017	0.137
$\sigma 2$	1.000	0.906	9.4	0.111	0.040	0.118	0.984	1.6	0.136	0.016	0.137
$\sigma 3$	1.000	0.867	13.3	0.112	0.041	0.119	0.941	5.9	0.135	0.017	0.136
$\sigma 4$	1.000	0.904	9.6	0.111	0.040	0.118	0.982	1.8	0.136	0.017	0.137
$\sigma 5$	1.000	0.927	7.3	0.117	0.041	0.124	1.002	0.2	0.140	0.016	0.141
Overall mean value across parameters		–	9.8%	0.121	0.045	0.129	–	2.5	0.151	0.019	0.153
Mean time		66.09					1.96				
Std. dev. of time		10.87					0.42				
% of Runs converged		100%					100%				

4.性能比較結果 MSL vs MACML(1)

Cross-sectional random coefficients model

【Non Diagonal case】

- 分散共分散行列の下三角を推定
- Diagonal caseと同様に約33倍早い（バイアスは2.1ポイント）

$$\tilde{\Omega} = \begin{bmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & & \\ l_{31} & l_{32} & l_{33} & & \\ l_{41} & l_{42} & l_{43} & l_{44} & \\ l_{51} & l_{52} & l_{53} & l_{54} & l_{55} \end{bmatrix}$$

Table 1b
Evaluation of the ability to recover true parameters for the cross−sectional non-diagonal case.

Parameter	True value	MSL method					MACML method				
		Parameter estimates		Standard error estimates			Parameter estimates		Standard error estimates		
		Mean estimate	Absolute percentage bias (%)	Asymptotic standard error	Simulation standard error	Simulation adjusted asymptotic standard error	Mean estimate	Absolute percentage bias (%)	Asymptotic standard error	Approximation standard error	Approximation adjusted asymptotic standard error
Mean values of the β vector											
*b*1	1.500	1.374	8.4	0.133	0.049	0.142	1.443	3.8	0.147	0.022	0.148
*b*2	−1.000	−0.912	8.8	0.093	0.037	0.100	−0.959	4.1	0.102	0.014	0.103
*b*3	2.000	1.830	8.5	0.174	0.068	0.187	1.923	3.8	0.191	0.029	0.193
*b*4	1.000	0.914	8.6	0.092	0.032	0.097	0.958	4.2	0.101	0.014	0.102
*b*5	−2.000	−1.849	7.6	0.176	0.068	0.189	−1.941	3.0	0.194	0.028	0.196
Cholesky parameters characterizing the covariance matrix of the β vector											
*l*11	1.000	0.909	9.1	0.112	0.040	0.119	0.959	4.1	0.119	0.017	0.120
*l*12	−0.500	−0.463	7.3	0.085	0.029	0.090	−0.472	5.6	0.085	0.010	0.085
*l*13	0.250	0.231	7.5	0.089	0.036	0.096	0.233	6.7	0.087	0.009	0.088
*l*14	0.750	0.689	8.2	0.092	0.028	0.097	0.707	5.7	0.095	0.013	0.096
*l*15	0.000	0.006	0.6	0.086	0.040	0.095	0.015	1.5	0.088	0.008	0.089
*l*22	0.866	0.756	12.7	0.109	0.043	0.117	0.809	6.5	0.116	0.017	0.117
*l*23	0.433	0.431	0.5	0.105	0.050	0.117	0.436	0.6	0.100	0.012	0.101
*l*24	−0.144	−0.149	3.6	0.101	0.041	0.109	−0.170	17.8	0.093	0.010	0.094
*l*25	0.000	−0.021	2.1	0.101	0.055	0.115	−0.019	1.9	0.098	0.010	0.099
*l*33	0.866	0.750	13.4	0.130	0.073	0.149	0.812	6.3	0.131	0.019	0.132
*l*34	0.237	0.242	2.0	0.112	0.055	0.125	0.259	9.3	0.106	0.011	0.106
*l*35	0.000	−0.031	3.1	0.120	0.081	0.145	−0.029	2.9	0.116	0.011	0.117
*l*44	0.601	0.464	22.9	0.126	0.085	0.152	0.531	11.6	0.125	0.015	0.126
*l*45	0.000	−0.053	5.3	0.168	0.134	0.214	−0.053	5.3	0.171	0.017	0.172
*l*55	1.000	0.885	11.5	0.125	0.089	0.153	0.956	4.4	0.136	0.018	0.137
Overall mean value across parameters		–	7.6	0.116	0.057	0.130	–	5.5	0.120	0.015	0.121
Mean time		174.32					5.19				
Std. dev. of time		28.13					0.84				
% of Runs converged		100					100				

5. まとめ

- 正規分布を持つOpen-formなモデル（Mixed-Probit）を対象とした新たな推定手法（MACML推定）を提案
- パネルや空間相関を考慮した改良モデルに対してもMACML推定が適用可能であることを提示
- 数値実験より高速かつ低バイアスなパラメータ推定値が得られる事を確認

【著者談】

In closing, the MACML inference approach has the potential to dramatically influence the use of the mixed multinomial probit model in practice, and should facilitate the practical application of rich model structures for unordered-response discrete choice modeling.