多項プロビットモデルの ベイズ推定への一考察

A Bayesian estimation of the multinomial probit model

宮　脇　幸　治

The multinomial probit model is a typical statistical model for a discrete choice analysis. To estimate its model parameters by taking the Bayesian approach, the sampling inefficiency is of interest. Recently, the marginal data augmentation framework has been proposed to address this issue. This study proposes a Bayesian estimation method for the multinomial probit model by extending this framework.

Koji Miyawaki

JEL：C11, C25
キーワード：離散選択、多項プロビットモデル
Keywords： Discrete choice, Multinomial probit model

I　はじめに

多項プロビットモデルとは，離散的な選択を行うような意思決定を統計的モデルとして記述したものである．例えば，通勤形態の選択を考える．徒歩，電車，バスの 3 形態あった場合の選択にどのような要因が関係しているかということを分析する際に，多項プロビットモデルを用いることがある．

統計的モデルの見地からは，多項プロビットモデルとは通常のプロビットモデルの拡張といえる．プロビットモデルとは，被説明変数が二値のみとる場合に用いられる統計的モデルである．先ほどの通勤形態の選択においては，形態が 3 つあるため，プロビットモデルを用いることはできない．このような場合，その自然な拡張としての多項プロビットモデルが用いられる．

被説明変数が二値のみとる場合の統計的モデルとしては，ロジットモデルもあげられる．しかし，このモデルにおいては無意味な選択肢からの独立を仮定する必要があり，この仮定が成立しない状況ではロジットモデルを用いることはできない．なおプロビットモデルにおいては，このような仮定を置く必要がない．

ベイズアプローチによる多項プロビットモデルの統計的推測には，プロビットモデルにおける Albert and Chib (1993) を多項に拡張した McCulloch and Rossi (1994) によるものがしばしば用いられる．これらの論文では，プロビットモデル及び多項プロビットモデルにおいて，直接パラメータの事後分布を解析したり，その標本を得ることが難しいため，マルコフ連鎖モンテカルロ法（MCMC）を用いて事後分布からの標本を近似的に得た上で，統計的推測を行う．統計的モデルについての詳細は後述するが，彼らの手法はモデルにおける潜在変数も含めて MCMC によるサンプリングを行うところが特徴である．

しかし，その手法は非効率的となることがある．つまり事後分布からの標本とみなすために，非常に多くの MCMC 標本を必要とすることがある．そのような非効率性は多項プロビットモデルにおいて顕著となる．

非効率性を改善するための手法はいくつか提案されている．まず，Maruyama and Strawderman (2012) では潜在変数を積分消去した事後分布を，無情報の事前分布を仮定したプロビットモデルにおいて導出している．しかしこれを多項の場合に拡張することは容易ではない．また，Liu and Sabatti (2000) で提案されている一般化ギブスサンプリングを応用することも考えられるが，あまり非効率性が改善しない場合も見うけられる．

一方 Imai and van Dyk (2005) では marginal data augmentation と呼ばれるアプローチによる推定手法が提案されている．この手法は，多項プロビットモデルにおける効率性の改善が高いことが知られているため，本論文ではこの手法を元にした多項プロビットモデルのベイズ推定方法を提案する．

II　多項プロビットモデル

まず多項プロビットモデルについての説明を行う．選択肢が $p+1$ 個あると

する．通勤形態の例では，$p=2$ である．選択された結果が y_i（$i=1,\ldots,n$）であり，0 から p までの整数値で表されているとする．また選択を説明する変数として $\boldsymbol{X}_i=(\boldsymbol{x}_{i1},\ldots,\boldsymbol{x}_{ip})'$ があるとする．それぞれ $\boldsymbol{x}_{ij}$（$j=1,\ldots,p$）は選択肢 j に対する k 次元のベクトルである．

このとき多項プロビットモデルは

$$y_i=j, \quad \text{if} \max_{j=0,\ldots,p} y_{ij}^*, \tag{1}$$

$$y_{ij}^*=\boldsymbol{x}_{ij}'\boldsymbol{\beta}+u_{ij}, \quad j=1,\ldots,p, \tag{2}$$

と表すことができる．ここで，y_{ij}^* はそれぞれの選択肢に対する観察されない評価（計量経済学の文脈では潜在効用と呼ばれる）で，$y_{i0}^*=0$ として正規化してある．もし正規化を行なっていなければ，どの選択肢への評価が最大であるかのみが条件であるため，すべての評価に定数を加えても条件を満たしてしまう．その場合，パラメータの識別ができないため，このような制約を課す．また，$\boldsymbol{\beta}$ は潜在効用の回帰係数，$\boldsymbol{u}_i=(u_{i1},\ldots,u_{ip})'$ は平均 $\mathbf{0}$，分散共分散行列 $\boldsymbol{\Sigma}$ の p 変量正規分布に従う誤差項であり $N(\mathbf{0},\boldsymbol{\Sigma})$ と表記する．

この統計的モデルが表していることは，$p+1$ 個の選択肢から一つ選ぶとき，それぞれの選択肢から得られるであろう効用と基準とする最初の選択肢からの効用の差を y_{ij}^*（$j\neq 0$）として，その中から最大の効用が得られるような選択肢を選ぶという意思決定プロセスである．そのような意思決定は消費者の効用最大化というミクロ経済学の基礎となる仮定とも整合的であるため，計量経済学の分野ではこのような統計的モデルが用いられることが多い．

また効用の大きさではなく順序に意味があるため，効用を定数倍しても選択されるものは変わらない．モデルはそのように作られているが，パラメータを推定するには制約を課す必要がある．多くの場合，y_{i1}^* の分散を 1，$\sigma_{11}=1$，とする制約を課す．

ここで説明されている制約については，どの選択肢に制約を課すかによって結果が異なることも指摘されている．そのような場合は，Burgette and Nordheim (2012) もしくは Burgette and Hahn (2013) で提案されているような手法を用いることが望ましい．しかし，簡単化のため，本論文ではこれらに関する議論は

行わない．

III　Imai and van Dyk (2005) の手法

分散共分散行列の (1,1) 要素が 1 である制約を考慮に入れた多項プロビットモデルの推定手法に関しては，これまで様々な方法が提案されてきた．最も簡単なものは，制約なしで推定し，事後的に制約を課すもので，McCulloch and Rossi (1994) によって提案された．また，制約を課した分散共分散行列をパラメータ変換し，コレスキー分解を用いる推定手法が McCulloch, Polson, and Rossi (2000) によって提案されている．これらの方法以外には，marginal data augmentation と呼ばれる一般的なフレームワークを多項プロビットモデルに応用した手法が，Imai and van Dyk (2005) によって提案されている．Marginal data augmentation 自体については，van Dyk and Meng (2001) などを参照されたい．本論文の目的は，彼らのアプローチを拡張し，効率的な推定手法を提案することである．そのため，以下では彼らのアプローチについて簡単に説明を行う．

まず定数倍に関する制約を以下のように緩める．パラメータ $\alpha\,(>0)$ を導入し，潜在効用，回帰のパラメータ及び誤差項を α 倍したモデルを考える．どの選択肢が選ばれるかは変わらないが，潜在効用のモデルは以下のように書き換えられる．

$$\tilde{y}_{ij} = \boldsymbol{x}'_{ij}\tilde{\boldsymbol{\beta}} + \tilde{u}_{ij}. \tag{3}$$

ここで $(\tilde{y}_{ij}, \tilde{\boldsymbol{\beta}}, \tilde{u}_{ij}) = (\alpha y^*_{ij}, \alpha\boldsymbol{\beta}, \alpha u_{ij})$ である．この時，誤差項の分散共分散行列は $\tilde{\boldsymbol{\Sigma}} = \alpha^2\boldsymbol{\Sigma}$ となる．このモデルにおいて，もし $\tilde{\boldsymbol{\Sigma}}$ を推定できれば，その (1,1) 要素として α^2 が推定されるので，その平方根を用いて元のモデルのパラメータを推定することができる．このようなパラメータ α をモデルのパラメータと区別して，ワーキングパラメータと呼ぶことがある．

III.1　ベイズアプローチ

Imai and van Dyk (2005) の手法はベイズ統計学の枠組みを用いているため，以下その枠組みを説明する．モデルのパラメータ $(\boldsymbol{\beta}, \tilde{\boldsymbol{\Sigma}})$ に，以下のような事

前分布を仮定する．

$$\boldsymbol{\beta} \sim N\left(\boldsymbol{b}_0, \boldsymbol{B}_0^{-1}\right), \tag{4}$$

$$\tilde{\boldsymbol{\Sigma}} \sim IW\left(\nu_0 + p, \tilde{\boldsymbol{S}}_0\right). \tag{5}$$

ここで，$\tilde{\boldsymbol{S}}_0 = \alpha_0^2 \boldsymbol{S}_0$ である．下付きの添え字の 0 は，事前分布のパラメータであり，既知であることを意味している．

分散共分散行列の事前分布において，$\boldsymbol{V} \sim IW(r, \boldsymbol{R})$ は p 次元の行列 V が逆ウィッシャート分布に従うことを表しており，その比例定数を除く確率密度は

$$p(\boldsymbol{V}) \propto |\boldsymbol{R}|^{(r-p-1)/2} |V|^{-r/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{R}\boldsymbol{V}^{-1}\right)\right\}, \tag{6}$$

である．逆ウィッシャート分布の確率密度関数については，いくつかパラメータの置き方が異なるものが用いられるが，本論文では Gupta and Nagar (2000) で定義されているものを用いている．この分布のさらなる性質についてもこちらを参考にされたい．

事前分布において，後者の $\tilde{\boldsymbol{\Sigma}}$ の事前分布を $(\boldsymbol{\Sigma}, \alpha^2)$ の分布として見た上で，分解すれば

$$\alpha^2 \mid \boldsymbol{\Sigma} \sim IG\left[\frac{\nu_0(p-1)}{2}, \frac{\alpha_0^2}{2}\left\{\mathrm{tr}\left(\boldsymbol{S}_0\boldsymbol{\Sigma}^{-1}\right)\right\}\right], \tag{7}$$

$$p(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{\nu_0+p}{2}}\left\{\mathrm{tr}\left(\boldsymbol{S}_0\boldsymbol{\Sigma}^{-1}\right)\right\}, \tag{8}$$

が得られる．ここで $x \sim IG(a,b)$ とは，x が逆ガンマ分布に従うことを表し，a と b はそれぞれ形状母数と尺度母数を表す．その確率密度関数は，比例部分を除いて

$$p(x) \propto x^{-(a+1)} \exp\left(-bx^{-1}\right), \tag{9}$$

で与えられる．また $p(\boldsymbol{\Sigma})$ は $\boldsymbol{\Sigma}$ の事前確率密度である．

ここまでの議論を下にすれば，多項プロビットモデルの事後確率密度関数は

$$\begin{aligned}\pi\left(\{\tilde{\boldsymbol{y}}_i\}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \alpha^2\right) \propto \prod_{i=1}^{n}\left|\alpha^2\boldsymbol{\Sigma}\right|^{-1/2} \exp\left\{-\frac{1}{2}\left(\tilde{\boldsymbol{y}}_i - \boldsymbol{X}_i\alpha\boldsymbol{\beta}\right)' \alpha^{-2}\boldsymbol{\Sigma}^{-1}\left(\tilde{\boldsymbol{y}}_i - \boldsymbol{X}_i\alpha\boldsymbol{\beta}\right)\right\} \\ \times I\left(y_i = \underset{j\in\{0,1,\ldots,p\}}{\arg\max}\ \tilde{y}_{ij}\right)\pi\left(\boldsymbol{\beta}, \boldsymbol{\Sigma}, \alpha^2\right),\end{aligned} \tag{10}$$

である．ただし $\tilde{y}_{i0} = 0$，$\tilde{\boldsymbol{y}}_i = (\tilde{y}_{i1}, \ldots, \tilde{y}_{ip})'$，$\pi(\boldsymbol{\beta}, \boldsymbol{\Sigma}, \alpha^2)$ は事前分布の確率密度を表す．

III.2　ギブスサンプラー

前節で得られた事後分布を用いてパラメータの統計的推測を行う．しかし，この分布を直接解析することは難しい．そのため，MCMC を用いて事後分布からの近似的な標本を得た上で，得られた標本を用いて統計的推測を行うことが通常である．Imai and van Dyk (2005) ではギブスサンプラーと呼ばれる MCMC のアルゴリズムを二種類提案している．一つは，回帰係数 $\boldsymbol{\beta}$ の事前平均がゼロベクトルと仮定される場合に用いられるアルゴリズムであり，もう一つはこの仮定が緩められているものである．以下，後者において問題となりうる点を指摘するために，後者のアルゴリズムを紹介する．

Step 1. パラメータ及び潜在効用に初期値を与える: $(\{\boldsymbol{y}_i^*\}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \alpha^2)$.

Step 2. $\boldsymbol{\Sigma}$ を与えた α^2 の条件付き事後分布（事前分布と同じ）から標本を発生する．

$$\alpha^2 \mid \boldsymbol{\Sigma} \sim IG\left\{\frac{\nu_0(p-1)}{2}, \frac{\alpha_0^2}{2}\mathrm{tr}\left(\boldsymbol{S}_0\boldsymbol{\Sigma}^{-1}\right)\right\}. \tag{11}$$

Step 3. $\boldsymbol{\beta}$ と $\boldsymbol{\Sigma}$ を与えた $\{\boldsymbol{y}_i^*\}_i$ の条件付き事後分布から標本を発生する（$i = 1, \ldots, n$）．

$$\boldsymbol{y}_i^* \mid \boldsymbol{\beta}, \boldsymbol{\Sigma} \sim TN_{C_i}(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma}), \quad C_i = I\left(y_i = \arg\max_j y_{ij}^*\right). \tag{12}$$

ここで $TC_A(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ は，台 A 上の平均ベクトル $\boldsymbol{\mu}$，分散共分散行列 $\boldsymbol{\Sigma}$ の切断多変量正規分布である．この分布からの最も簡単な標本の発生方法は，j 番目の要素以外の要素を与えた y_{ij}^* の条件付き事後分布（一次元の切断正規分布）から標本を発生させることを $j = 1, \ldots, p$ について順に行うことである．

Step 4. $\tilde{u}_{ij} = \alpha(y_{ij}^* - \boldsymbol{x}_{ij}'\boldsymbol{\beta})$ とする．

Step 5. $\{\tilde{\boldsymbol{u}}_i\}_{i=1}^n$ を与えた $(\boldsymbol{\Sigma}, \alpha^2)$ の条件付き事後分布から標本を発生する．これは $\tilde{\boldsymbol{\Sigma}}$ からの標本発生と同一視できる．

$$\tilde{\boldsymbol{\Sigma}} \mid \{\tilde{\boldsymbol{u}}_i\}_{i=1}^n \sim IW\left(\nu_0 + p + n, \tilde{\boldsymbol{S}}_0 + \sum_{i=1}^n \tilde{\boldsymbol{u}}_i\tilde{\boldsymbol{u}}_i'\right). \tag{13}$$

ここで $\tilde{\boldsymbol{u}}_i = (\tilde{u}_{i1}, \ldots, \tilde{u}_{ip})'$ である．

Step 6. $y_{ij}^* = \tilde{u}_{ij}/\sqrt{\tilde{\sigma}_{11}} + \boldsymbol{x}_{ij}'\boldsymbol{\beta}$ とする．

Step 7. Step 3-6 を，$y_i = \arg\max_j y_{ij}^*$ となるまで繰り返す．

Step 8. $\{\boldsymbol{y}_i^*\}_{i=1}^n$ と $\boldsymbol{\Sigma}$ を与えた $\boldsymbol{\beta}$ の条件付き事後分布から標本を発生する．

$$\boldsymbol{\beta} \mid \{\boldsymbol{y}_i^*\}_{i=1}^n, \boldsymbol{\Sigma} \sim N\left(\boldsymbol{b}_1, \boldsymbol{B}_1^{-1}\right), \tag{14}$$

ただし

$$\boldsymbol{B}_1^{-1} = \boldsymbol{B}_0 + \sum_{i=1}^n \boldsymbol{X}_i'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_i, \tag{15}$$

$$\boldsymbol{b}_1 = \boldsymbol{B}_1\left(\boldsymbol{B}_0\boldsymbol{b}_0 + \sum_{i=1}^n \boldsymbol{X}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{y}_i^*\right). \tag{16}$$

Step 9. Step 2-8 を十分長く繰り返す．

IV　Imai and van Dyk (2005) の拡張

IV.1　提案するアルゴリズム

先述した Imai and van Dyk (2005) のギブスサンプラーは，データによっては計算時間が長くかかる可能性が残されている．これは Step 7 において，多項プロビットモデルの制約条件（いわゆる効用最大化条件）を満たすまで繰り返す部分が含まれているからである．本論文では，彼らの前者のアルゴリズムを拡張し，棄却法を用いることで，この問題の改善を試みる．まず最初の 3 ステップは前節のアルゴリズムと同じであるため，詳細は省略する．

Step 1. パラメータ及び潜在効用に初期値を与える: $(\{\boldsymbol{y}_i^*\}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \alpha^2)$.

Step 2. $\boldsymbol{\Sigma}$ を与えた α^2 の条件付き事後分布（事前分布と同じ）から標本を発生する．

Step 3. $\boldsymbol{\beta}$ と $\boldsymbol{\Sigma}$ を与えた $\{y_i^*\}_{i=1}^n$ の条件付き事後分布から標本を発生する．

Step 4. $\tilde{y}_{ij} = \alpha y_{ij}^*$ とする．

Step 5. $\{\tilde{\boldsymbol{y}}_i\}_{i=1}^n$ と $\boldsymbol{\Sigma}$ を与えた α^2 の条件付き事後分布から標本を発生する．この条件付き事後確率密度は

$$p\left(\alpha^2 \mid \{\tilde{\boldsymbol{y}}_i\}_{i=1}^n, \boldsymbol{\Sigma}\right) \propto \left(\alpha^2\right)^{-(a_0/2+1)} \exp\left\{-\frac{1}{2}\left(a_1\alpha^{-2} - 2a_2\alpha^{-1}\right)\right\}, \tag{17}$$

ただし $a_0 = (\nu_0 + n)(p-1)$,

$$a_1 = \alpha_0^2 \mathrm{tr}\left(\boldsymbol{S}_0\boldsymbol{\Sigma}^{-1}\right) + \sum_{i=1}^n \tilde{\boldsymbol{y}}_i'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{y}}_i - \boldsymbol{m}_a'\left(\boldsymbol{B}_0 + \boldsymbol{S}_a\right)^{-1}\boldsymbol{m}_a, \tag{18}$$

$$a_2 = \boldsymbol{b}_0'\boldsymbol{B}_0\left(\boldsymbol{B}_0 + \boldsymbol{S}_a\right)^{-1}\boldsymbol{m}_a, \tag{19}$$

$$\boldsymbol{m}_a = \sum_{i=1}^n \boldsymbol{X}_i'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{y}}_i, \tag{20}$$

$$\boldsymbol{S}_a = \sum_{i=1}^n \boldsymbol{X}_i'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_i, \tag{21}$$

で与えられる．この条件付き事後密度は非標準であるため，直接この確率密度が表す分布から標本を発生することはできない．ただし $\boldsymbol{b}_0 = \boldsymbol{0}$ とすれば $a_2 = 0$ となり，条件付き事後分布は逆ガンマ分布となる．これは，Imai and van Dyk (2005) において $\boldsymbol{b}_0 = \boldsymbol{0}$ と仮定できるときに提案されているものである．このステップにおける標本発生方法は後述する．

Step 6. $\{\tilde{\boldsymbol{y}}_i\}_{i=1}^n$ 及び $\boldsymbol{\Sigma}$, α^2 を与えた $\tilde{\boldsymbol{\beta}}$ の条件付き事後分布から標本を発生する．

$$\tilde{\boldsymbol{\beta}} \mid \{\tilde{\boldsymbol{y}}_i\}_{i=1}^n, \boldsymbol{\Sigma}, \alpha^2 \sim N\left\{\alpha\left(\boldsymbol{B}_0 + \boldsymbol{S}_a\right)^{-1}\left(\boldsymbol{B}_0\boldsymbol{b}_0 + \alpha^{-1}\boldsymbol{m}_a\right), \alpha^2\left(\boldsymbol{B}_0 + \boldsymbol{S}_a\right)^{-1}\right\}. \tag{22}$$

Step 7. $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}/\alpha$ とする．

Step 8. $\{\tilde{\boldsymbol{y}}_i\}_{i=1}^n$ と $\tilde{\boldsymbol{\beta}}$ を与えた $\tilde{\boldsymbol{\Sigma}}$ の条件付き事後分布から標本を発生する．

$$\tilde{\boldsymbol{\Sigma}} \mid \{\tilde{\boldsymbol{y}}_i\}_{i=1}^n, \tilde{\boldsymbol{\beta}} \sim IW\left\{\nu_0 + p + n, \tilde{\boldsymbol{S}}_0 + \sum_{i=1}^n \left(\tilde{\boldsymbol{y}}_i - \boldsymbol{X}_i\tilde{\boldsymbol{\beta}}\right)\left(\tilde{\boldsymbol{y}}_i - \boldsymbol{X}_i\tilde{\boldsymbol{\beta}}\right)'\right\}. \tag{23}$$

Step 9. $\boldsymbol{\Sigma} = \tilde{\boldsymbol{\Sigma}}/\tilde{\sigma}_{11}$, $\boldsymbol{y}_i^* = \tilde{\boldsymbol{y}}_i/\sqrt{\tilde{\sigma}_{11}}$ とする．

Step 10. Step 2-9 を十分長く繰り返す．

IV.2　Step 5

前述の通り，α^2 の条件付き事後確率密度は非標準のものであるため，直接標本を発生することが難しい．そのため，本論文ではこの確率密度関数を近似する包絡関数を求め，棄却法を用いて標本を発生させる方法を提案する．棄却法については，例えば古澄 (2015) を参照されたい．以下 a_2 の符号に従って場合分けが必要となるため，それぞれのアルゴリズムを順に説明する．

$a_2 \le 0$ の場合

Step 5-1 候補 $\alpha^\star$ を以下の逆ガンマ分布より発生させる.

$$(\alpha^\star)^2 \sim IG\left(\frac{a_0}{2}, \frac{a_1}{2}\right). \tag{24}$$

Step 5-2 $u \sim U(0,1)$ から発生させる.

Step 5-3 もし

$$u \le \exp\left(a_2/\alpha^\star\right), \tag{25}$$

なら，候補 $\alpha^\star$ を受容する．そうでなければ棄却する.

$a_2 > 0$ の場合

Step 5-1 候補 $\alpha^\star$ を以下の逆ガンマ分布より発生させる.

$$\alpha^\star \sim IG\left(a_0, a_2\right). \tag{26}$$

Step 5-2 $u \sim U(0,1)$ から発生させる.

Step 5-3 もし

$$u \le \exp\left\{-\frac{1}{2}a_1\left(\alpha^\star\right)^{-2} + 2a_2\left(\alpha^\star\right)^{-1} - \frac{2a_2^2}{a_1}\right\}, \tag{27}$$

なら，候補 $\alpha^\star$ を受容する．そうでなければ棄却する.

いずれのアルゴリズムにおいても，条件付き事後密度と候補を発生させる分布の密度の比は有限な定数を上限として持つことは明らかである．また，前者のアルゴリズムについてコメントを 2 つ述べる．まず，$a_2 = 0$ であれば，常に候補を受容していることに注意する．また $a_2 \le 0$ において，逆ガンマ分布が正しく定義されるには $a_1 > 0$ である必要がある．以下はその証明である.

Proposition 1. 式 (18) で定義されている a_1 は正.

Proof. まず $\boldsymbol{z}_i = \boldsymbol{\Sigma}^{-1/2}\tilde{\boldsymbol{y}}_i$，$\boldsymbol{A}_i = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{X}_i$ とする（$i = 1, \ldots, n$）．このとき a_1 を構成する最後の二項は

$$\sum_{i=1}^n \tilde{\boldsymbol{y}}_i'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{y}}_i - \boldsymbol{m}_a'\left(\boldsymbol{B}_0 + \boldsymbol{S}_a\right)^{-1}\boldsymbol{m}_a = \sum_{i=1}^n \boldsymbol{z}_i'\boldsymbol{z}_i - \left(\sum_{i=1}^n \boldsymbol{A}_i'\boldsymbol{z}_i\right)'\left(\boldsymbol{B}_0 + \sum_{i=1}^n \boldsymbol{A}_i'\boldsymbol{A}_i\right)^{-1}\left(\sum_{i=1}^n \boldsymbol{A}_i'\boldsymbol{z}_i\right), \tag{28}$$

と変形できる．ここで $\boldsymbol{z} = (\boldsymbol{z}_1', \ldots, \boldsymbol{z}_n')'$，$\boldsymbol{A} = (\boldsymbol{A}_1', \ldots, \boldsymbol{A}_n')'$ とすれば

$$\boldsymbol{z}' \left\{ \boldsymbol{I} - \boldsymbol{A} \left(\boldsymbol{B}_0 + \boldsymbol{A}'\boldsymbol{A} \right)^{-1} \boldsymbol{A}' \right\} \boldsymbol{z}, \tag{29}$$

となる．

さて，分散共分散行列 $\boldsymbol{B}_0 > \boldsymbol{O}$，$\boldsymbol{\Sigma} > \boldsymbol{O}$ であり，$\boldsymbol{X}_i$ のランクが列数に等しく，$\tilde{\boldsymbol{y}}_i = \boldsymbol{0}$ となることが確率 0 であることから，Bergstrom の不等式より，

$$\boldsymbol{z}'\boldsymbol{A} \left(\boldsymbol{B}_0 + \boldsymbol{A}'\boldsymbol{A} \right)^{-1} \boldsymbol{A}'\boldsymbol{z} \leq \frac{\left(\boldsymbol{z}'\boldsymbol{A}\boldsymbol{B}_0^{-1}\boldsymbol{A}'\boldsymbol{z} \right) \left\{ \boldsymbol{z}'\boldsymbol{A} \left(\boldsymbol{A}'\boldsymbol{A} \right)^{-1} \boldsymbol{A}'\boldsymbol{z} \right\}}{\boldsymbol{z}'\boldsymbol{A} \left\{ \boldsymbol{B}_0^{-1} + \left(\boldsymbol{A}'\boldsymbol{A} \right)^{-1} \right\} \boldsymbol{A}'\boldsymbol{z}} < \boldsymbol{z}'\boldsymbol{A} \left(\boldsymbol{A}'\boldsymbol{A} \right)^{-1} \boldsymbol{A}'\boldsymbol{z}. \tag{30}$$

最後の不等号は，$a, b > 0$ に対して $a/(a+b) < 1$ を用いた．また Bergstrom の不等式については Abadir and Magnus (2005) の Exercise 12.3（323 ページ）を参照のこと．

よって $a_1 > 0$ には，

$$\boldsymbol{z}' \left\{ \boldsymbol{I} - \boldsymbol{A} \left(\boldsymbol{A}'\boldsymbol{A} \right)^{-1} \boldsymbol{A}' \right\} \boldsymbol{z} \geq 0, \tag{31}$$

を示せば十分．これは対称かつべき等な行列が半正値定符号であることから示すことができる．□

V　おわりに

本論文では多項プロビットモデルに関する推定手法の改善について議論を行なった．特に marginal data augmentation と呼ばれるアプローチに基づく Imai and van Dyk (2005) の手法の改善を試みた．今回は推定手法の提案に留まっているが，実際のデータを用いた棄却法の評価が今後の課題となろう．

（筆者は関西学院大学経済学部准教授）

謝辞

本研究の一部は北原稔氏（大阪市立大学）との議論によるものである．ここに感謝の意を表する．

参考文献

Abadir, K. M. and J. R. Magnus (2005). *Matrix Algebra*, Volume 1 of *Econometric Exercises*. Cambridge, New York: Cambridge University Press.

Albert, J. H. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association 88*(422), 669–679.

Burgette, L. F. and E. V. Nordheim (2012). The trace restriction: an alternative identification strategy for the Bayesian multinomial probit model. *Journal of Business & Economic Statistics 30*(3), 404–410.

Burgette, L. F. and P. R. Hahn (2013). A symmetric prior for multinomial probit models. Working paper.

Gupta, A. K. and D. K. Nagar (2000). *Matrix Variate Distributions*, Volume 104 of *Monographs and Surveys in Pure and Applied Mathematics*. Boca Raton: Chapman & Hall/CRC.

Imai, K. and D. A. van Dyk (2005). A Bayesian analysis of the multinomial probit model using marginal data augmentation. *Journal of Econometrics 124*(2), 311–334.

Liu, J. S. and C. Sabatti (2000). Generalised Gibbs sampler and multigrid Monte Carlo for Bayesian computation. *Biometrika 87*(2), 353–369.

Maruyama, Y. and W. E. Strawderman (2012). A new Monte Carlo sampling in Bayesian probit regression. arXiv:1202.4339.

McCulloch, R. and P. E. Rossi (1994). An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics 64*(1-2), 207–240.

McCulloch, R. E., N. G. Polson, and P. E. Rossi (2000). A bayesian analysis of the multinomial probit model with fully identified parameters. *Journal of Econometrics 99*(1), 173–193.

The art of data augmentation. *Journal of Computational and Graphical Statistics 10*(1), 1–50.

古澄英男 (2015) 『ベイズ計算統計学』 朝倉書店.