多値ＩＲＴの条件付き最尤推定値の順序保存について

東京大学　総合文化研究科　森　一将

1. 概要とモデル

本研究では多値型 IRT モデルの 1 つである Rating Scale Model(RSM; Andrich,1978) と関連する多値 IRT モデルの母数の条件付き最尤推定値と順序性が保存する統計量を導出する．k 個の項目がそれぞれ m カテゴリを持ち,n 人の受験者から成るテストデータを考える．このとき,i 番目の受験者が j 番目の項目の h 番目のカテゴリに対する反応を $X_{ijh}=\{0,1\}(i=1,\cdots,n,j=1,\cdots,k,h=\cdots,m)$ とする．このとき,RSM によって対応する確率は以下のようになる．

$$P_{ijh}(\theta_i,\alpha_j)=P(X_{ijh}=1;\boldsymbol{\theta},\boldsymbol{\alpha})=\frac{\exp(w_h\theta_i+a_{jh})}{\sum_{h=1}^{m}\exp(w_h\theta_i+\alpha_{jh})}. \tag{1}$$

ここで,$\boldsymbol{\theta}=(\theta_1,\cdots,\theta_n)$ は受験者の能力母数を示し,$\boldsymbol{\alpha}=(\alpha_{11},\cdots,\alpha_{1m},\cdots,\alpha_{k1},\cdots,\alpha_{km})$ はテストの項目母数を示す．また,$w_h(h=1,\cdots,m)$ は重みづけ係数 (配点）であり, 所与とする．また, このモデルの拡張として, 所与の重みづけ係数 v_{jhp} と母数 $\eta_p(p=1,\cdots,q)$ を用いて $\alpha_{jh}=\sum_{p=1}^{q}v_{jhp}\eta_p$ と再母数化すると linear rating scale model(LRSM; Fischer and Parzer, 1991) と呼ばれるモデルとなり,$w_h=h$ かつ $\alpha_{jh}=\beta_{jh}$ とおくと,partial credit model (PCM; Masters, 1982) となり,$w_h=h$ かつ $\beta_{jh}=\sum_{p=1}^{q}u_{jhp}\gamma_p$（$u_{jhp}$ は所与の重みづけ係数,γ_p は母数）と再母数化すると linear partial credit model (LPCM; Glas and Verhelst,1989) となる．

まず,(1) を基にして項目母数の条件付き最尤推定値を考える．この時の条件付き尤度関数は以下のようになる．

$$L_C(\boldsymbol{\alpha}|\boldsymbol{t})=\frac{\exp(\sum_{j=1}^{k}\sum_{h=1}^{m}\alpha_{jh}\sum_{i=1}^{n}x_{ijh})}{\prod_{i=1}^{n}G_{t_i}(\boldsymbol{\alpha})}=\frac{\exp(\sum_{j=1}^{k}\sum_{h=1}^{m}\alpha_{jh}r_{jh})}{\prod_{t=1}^{km}[G_t(\boldsymbol{\alpha})]^{N_t}}.$$

ここで,$t_i=\sum_{j=1}^{k}\sum_{h=1}^{m}w_hx_{ijh}$ は i 番目の受験者の得点を示し,$r_{jh}=\sum_{i=1}^{n}x_{ijh}$ は j 番目の項目の h 番目のカテゴリに反応した受験者数を示す．また,$G_{t_i}(\boldsymbol{\alpha})=\sum_{(t_i)}\exp\alpha_{jh}x_{ijh}$,$\sum_{(t_i)}$ は t_i という得点の受験者に対応する反応パターンをすべて足し合わせる演算を表し, N_t は 得点 t をとった受験者の人数を示す．この尤度関数の一部に以下のような近似を考える．

$$G_t(\boldsymbol{\alpha})=\sum_{(t_i)}\exp\alpha_{jh}x_{ijh}\simeq m\prod_{(t_i)}\exp\frac{1}{m}\left\{\alpha_{jh}x_{ijh}\right\}=\tilde{G}_{t_i}(\boldsymbol{\alpha}).$$

すると, 条件付き尤度関数は以下のようになる．

$$L_C(\boldsymbol{\alpha}|\boldsymbol{t})\simeq\frac{\exp(\sum_{j=1}^{k}\sum_{h=1}^{m}\alpha_{jh}r_{jh})}{\prod_{t=1}^{km}[\tilde{G}_t(\boldsymbol{\alpha})]^{N_t}}=\tilde{L}_C(\boldsymbol{\alpha}|\boldsymbol{t}).$$

同様に,LRSM,PCM,LPCM についても条件付き尤度関数の近似 $\tilde{L}_C(\boldsymbol{\eta}|\boldsymbol{t})$,$\tilde{L}_C(\boldsymbol{\beta}|\boldsymbol{t})$,$\tilde{L}_C(\boldsymbol{\gamma}|\boldsymbol{t})$ も導出ができる．この近似は「相加平均 $\geq$ 相乗平均」の関係を応用したものであり,Spect(1960) によって近似をした場合の比の母数に関する上限が明らかになっている．つまり, これらの近似による条件付き尤度関数の近似誤差は一定の範囲にとどまる．本研究ではこれらの近似を用いた条件付き尤度関数について, 推定値の順序保存の性質を考えていく．

2. 主要な結果

主要な結果は以下の通り．

命題 1. $\hat{\boldsymbol{\alpha}}_C$ を RSM の対数条件付き尤度関数 $\log\tilde{L}_C(\boldsymbol{\alpha}|\boldsymbol{t},\boldsymbol{r})$ を最大化する推定値とする．このとき, $\hat{\boldsymbol{\alpha}}_C^*$ が $\log\tilde{L}_C(\boldsymbol{\alpha}|\boldsymbol{t},\boldsymbol{r}_\uparrow)$ を最大化する推定値となる必要十分条件は $\hat{\boldsymbol{\alpha}}_C^*=\hat{\boldsymbol{\alpha}}_{C\uparrow}$.

ここで,$\boldsymbol{t}=(t_1\cdots,t_n)$,$\boldsymbol{r}=(r_{11}\cdots,r_{1m},\cdots,r_{k1},\cdots,\cdots,r_{km})$ とし,$\boldsymbol{r}_\uparrow$ はベクトル $\boldsymbol{r}$ の昇順での並べ替えを示す．

命題 2. $\hat{\boldsymbol{\eta}}_C$ を LRSM の対数条件付き尤度関数 $\log\tilde{L}_C(\boldsymbol{\eta}|\boldsymbol{t},\boldsymbol{r}')$ を最大化する推定値とする．このとき, $\hat{\boldsymbol{\eta}}_C^*$ が $\log\tilde{L}_C(\boldsymbol{\eta}|\boldsymbol{t},\boldsymbol{r}'_\uparrow)$ を最大化する推定値となる必要十分条件は $\hat{\boldsymbol{\eta}}_C^*=\hat{\boldsymbol{\eta}}_{C\uparrow}$. ここで,$\boldsymbol{r}'=(r'_1\cdots,r'_p)$,$r'_p=\sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{h=1}^{m}v_{jhp}x_{ijh}$ とする．

系 3. $\hat{\boldsymbol{\beta}}_C$ を PCM の対数条件付き尤度関数 $\log\tilde{L}_C(\boldsymbol{\beta}|\boldsymbol{t},\boldsymbol{r})$ を最大化する推定値とする．このとき, $\hat{\boldsymbol{\beta}}_C^*$ が $\log\tilde{L}_C(\boldsymbol{\beta}|\boldsymbol{t},\boldsymbol{r}_\uparrow)$ を最大化する推定値となる必要十分条件は $\hat{\boldsymbol{\beta}}_C^*=\hat{\boldsymbol{\beta}}_{C\uparrow}$.

系 4. $\hat{\boldsymbol{\gamma}}_C$ を LRSM の対数条件付き尤度関数 $\log\tilde{L}_C(\boldsymbol{\eta}|\boldsymbol{t},\boldsymbol{r}^*)$ を最大化する推定値とする．このとき, $\hat{\boldsymbol{\gamma}}_C^*$ が $\log\tilde{L}_C(\boldsymbol{\gamma}|\boldsymbol{t},\boldsymbol{r}^*_\uparrow)$ を最大化する推定値となる必要十分条件は $\hat{\boldsymbol{\gamma}}_C^*=\hat{\boldsymbol{\gamma}}_{C\uparrow}$. ここで,$\boldsymbol{r}^*=(r^*_1\cdots,r^*_p)$,$r^*_p=\sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{h=1}^{m}u_{jhp}x_{ijh}$ とする．

証明及び関連する結果については当日示す.

主要参考文献

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, **43**, 561-573.

Masters,G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, vol. **47**, 149-174.

Specht, W. (1960). Zur theorie der elementaren Mittel. *Math. Z.*, **74**,91-98.