以下の内容は私のゼミでRの使用法を説明するために作成した資料です。ゼミホームページ上で公開しており、自由に参照し、2次利用して頂いて構いません。ただし、内容について、一応検証してありますが、間違いがあるかもしれません。間違いがあった場合でもそれによって生じるいかなる損害、不利益について責任は負いかねますますのでご了承下さい。

（ゼミ生へ）
ゼミのページに以下のRのプログラムをおいておくので、必要に応じてそれらを用いて論文を書いて下さい。また不明な点、あるいは間違った点を見つけた場合は私まで教えて下さい。より使いやすくなるようにプログラムの改良をしてくれのも大歓迎です。

パネル2項ロジットモデルのバイアス推定

固定効果がない場合

観測されるデータ

$$y_{it} = \begin{cases} 1 & \text{with probability } p_{it} \\ 0 & \text{with probability } 1-p_{it} \end{cases}, \quad i = 1,\ldots,N, \quad t = 1,\ldots,T$$

例えば、確率的効用モデルで個人 i の t 時点での選択肢1からの効用は K 行1列の説明変数ベクトル $\mathbf{x}_{it}$ を用いて

$$U_{it}^{(1)} = \mathbf{x}_{it}'\boldsymbol{\beta}_1 + \eta_{it}^{(1)}$$

と表せ、選択肢2からの効用は

$$U_{it}^{(0)} = \mathbf{x}_{it}'\boldsymbol{\beta}_0 + \eta_{it}^{(0)}$$

と表せるとする。選択肢1をとったら $y_{it} = 1$, 選択肢2をとったら $y_{it} = 0$ をとるとすると、個人 i が t 時点において選択肢1をとる確率は

$$\begin{aligned} p_{it} = \Pr(y_{it} = 1) &= \Pr(\mathbf{x}_{it}'\boldsymbol{\beta}_0 + \eta_{it}^{(0)} < \mathbf{x}_{it}'\boldsymbol{\beta}_1 + \eta_{it}^{(1)}) \\ &= \Pr(\eta_{it}^{(0)} - \eta_{it}^{(1)} < \mathbf{x}_{it}'(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_0)) \\ &= \Pr(\varepsilon_{it} < \mathbf{x}_{it}'\boldsymbol{\gamma}) \end{aligned}$$

と表せる。ここで $\varepsilon_{it} = \eta_{it}^{(0)} - \eta_{it}^{(1)}$, $\boldsymbol{\gamma} = \boldsymbol{\beta}_1 - \boldsymbol{\beta}_0$ とする。もしここで ε_{it}, i =1,…,N, t =1,…,T について iidのロジスティック分布を仮定した場合、得られるモデルはロジットモデルとよばれ、p_{it} はロジスティック分布の分布関数

$$\Pr(\varepsilon_{it} < x) = \Lambda(x) = \frac{\exp(x)}{1+\exp(x)}$$

より

$$p_{it} = \Pr(\varepsilon_{it} < \mathbf{x}'_{it}\boldsymbol{\gamma}) = \frac{\exp(\mathbf{x}'_{it}\boldsymbol{\gamma})}{1+\exp(\mathbf{x}'_{it}\boldsymbol{\gamma})}$$

となり、尤度関数は

$$L(\boldsymbol{\gamma}) = \prod_{i=1}^{N}\prod_{t=1}^{T} p_{it}^{y_{it}}(1-p_{it})^{1-y_{it}} \text{ 。}$$

対数尤度関数は

$$\begin{aligned}\log L(\boldsymbol{\gamma}) &= \sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}\log p_{it} + \sum_{i=1}^{N}\sum_{t=1}^{T}(1-y_{it})\log(1-p_{it}) \\ &= \sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}\{\mathbf{x}'_{it}\boldsymbol{\gamma} - \log[1+\exp(\mathbf{x}'_{it}\boldsymbol{\gamma})]\} - \sum_{i=1}^{N}\sum_{t=1}^{T}(1-y_{it})\log[1+\exp(\mathbf{x}'_{it}\boldsymbol{\gamma})] \\ &= \sum_{i=1}^{N}\sum_{t=1}^{T}\{y_{it}\mathbf{x}'_{it}\boldsymbol{\gamma} - \log[1+\exp(\mathbf{x}'_{it}\boldsymbol{\gamma})]\}\end{aligned}$$

となる。これを $\boldsymbol{\gamma}$ について最大化するパラメーターを求めればよい。これはただ iid データが NT 個あるだけなので通常の(クロスセクション方向だけの)2 項ロジットモデルと本質的には何も変わらない。

固定効果がある場合

今、選択肢 1 と 2 からの効用に固定効果があり、それぞれ、

$$U_{it}^{(1)} = \alpha_i^{(1)} + \mathbf{x}'_{it}\boldsymbol{\beta}_1 + \eta_{it}^{(1)} \quad \text{および} \quad U_{it}^{(0)} = \alpha_i^{(0)} + \mathbf{x}'_{it}\boldsymbol{\beta}_0 + \eta_{it}^{(0)}$$

としよう。この時、先ほどと同じ議論により

$$p_{it} = \Pr(\varepsilon_{it} < \delta_i + \mathbf{x}'_{it}\boldsymbol{\gamma})$$

となる。ここで $\delta_i = \alpha_i^{(1)} - \alpha_i^{(0)}$ である。この時、$\boldsymbol{\delta} = [\delta_1,\ldots,\delta_N]'$ も未知パラメーターとみなした場合、

対数尤度関数は

$$\log L(\boldsymbol{\gamma},\boldsymbol{\delta}) = \sum_{i=1}^{N}\sum_{t=1}^{T}\{y_{it}(\delta_i + \mathbf{x}'_{it}\boldsymbol{\gamma}) - \log[1+\exp(\delta_i + \mathbf{x}'_{it}\boldsymbol{\gamma})]\} \qquad (1)$$

となる。この対数尤度を最大化する $\boldsymbol{\gamma}$ を $\hat{\boldsymbol{\gamma}}$ とする。

ジャックナイフ法によってパラメーターベクトル $\boldsymbol{\gamma}$ を推定する。$s(=1,..,T)$期のデータを抜いた対数尤度を

$$\log L^{(s)}(\boldsymbol{\gamma},\boldsymbol{\delta}) = \sum_{i=1}^{N}\sum_{t=1,t\neq s}^{T}\{y_{it}(\delta_i + \mathbf{x}'_{it}\boldsymbol{\gamma}) - \log[1+\exp(\delta_i + \mathbf{x}'_{it}\boldsymbol{\gamma})]\}$$

とし、この対数尤度を最大化する$\boldsymbol{\gamma}$を$\hat{\boldsymbol{\gamma}}^{(s)}$とする。この時バイアス補正された$\boldsymbol{\gamma}$の最尤推定値は

$$\tilde{\boldsymbol{\gamma}} = T\hat{\boldsymbol{\gamma}} - (T-1)\sum_{s=1}^{T}\frac{\hat{\boldsymbol{\gamma}}^{(s)}}{T}$$

と定義されます。$\boldsymbol{\delta}$の推定量はこの$\hat{\boldsymbol{\gamma}}^{(s)}$を用いて

$$\log L(\boldsymbol{\delta}\,|\,\tilde{\boldsymbol{\gamma}}) = \sum_{i=1}^{N}\sum_{t=1}^{T}\{y_{it}(\delta_i + \mathbf{x}'_{it}\tilde{\boldsymbol{\gamma}}) - \log[1+\exp(\delta_i + \mathbf{x}'_{it}\tilde{\boldsymbol{\gamma}})]\}$$

を最大化する$\boldsymbol{\delta}$の値として定義できる。

これをRでプログラムしてみよう。y_{it} と$\mathbf{x}_{it}$ と未知パラメーターγと$\boldsymbol{\delta}$は次のようなNT行$(K+1)$列の行列と$(K+N)$行1列のベクトルでまとめられているとする。

$$\mathbf{Y} = \begin{bmatrix} y_{11} & \mathbf{x}'_{11} \\ \vdots & \vdots \\ y_{1T} & \mathbf{x}'_{1T} \\ y_{21} & \mathbf{x}'_{21} \\ \vdots & \vdots \\ y_{2T} & \mathbf{x}'_{2T} \\ \vdots & \vdots \\ y_{N1} & \mathbf{x}'_{N1} \\ \vdots & \vdots \\ y_{NT} & \mathbf{x}'_{NT} \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\delta} \end{bmatrix} \qquad (2)$$

まず$\hat{\gamma}$を求めるRプログラムは(例えば)以下のようになる（なお以下の関数は`panelBLest.R`というRファイルとしてゼミのページにおいてあります）。まず`panelBL`を対数尤度関数を計算する関数として

```
panelBL=function(Y,N){
 T=nrow(Y)/N; K=ncol(Y)-1;
 X=matrix(Y[,2:(K+1)],nrow=N*T);
 y=matrix(Y[,1],nrow=N*T);
 function(S){
 Gam=matrix(S[1:K],nrow=K);
 Del=matrix(S[(K+1):(K+N)],nrow=N);
 iT = matrix(rep(1,T),nrow=T);

 Z=Del%x%iT + X%*%Gam;
 f=sum(y*Z - log(1+exp(Z)));
```

```
 y=-f; return(y)}
}

# Y: data matrix
# T: number of time periods
# N: number of individuals
# K: number of explanatory variables
```

とする。次に最尤法によって推定するプログラムを

```
panelBLMLE=function(G0,Y,N){
 Y=as.matrix(Y); K=ncol(Y)-1;
 S0=rbind(matrix(G0,ncol=1),matrix(rep(0,N),ncol=1));
 Shat=optim(S0,panelBL(Y,N),gr=NULL,method=c("BFGS"),
 control=list(maxit=100000,ndeps=(1e-5*rep(1,(K+N))),
 reltol=1e-10));
 Ghat=matrix(Shat$par[1:K],K,1)
 return(Ghat)
}
```

とする。`panelBL` は対数尤度を求める関数、`panelBLMLE` は先ほどの $\hat{\gamma}$ を求める関数である。$\hat{\gamma}$ を計算するための初期値を `G0` とすると、

```
Ghat=panelBLMLE(G0,Y,N)
```

とすれば `Ghat` が $\hat{\gamma}$ である。このプログラムがうまく動いてくれるか見るために、適当に発生させたデータに対して走らせてみよう。y_{it} として、N=10, T=5 としたものおよび $\mathbf{x}_{it}$ を K×1 ベクトル（K は説明変数の数）としたものとして $\mathbf{Y}$ を発生させる。

```
N=10; T=5; K=3;
y=matrix(0,nrow=N*T,ncol=1);
x=matrix(rnorm(N*T*K,0,1),ncol=K);
for (i in 1:N){
 y[(1+T*(i-1)):(T*i)]=matrix(rbinom(T,1,((N+i)/(3*N))),
 nrow=T,ncol=1)
}
Y=cbind(y,x);
```

(これは関数にしてもよいかもしれない)。ここで説明変数は標準正規分布に従う確率変数、被説明変数は $y_{it} \sim B((N+i)/3N)$ とした。$B(p)$ は確率 p で 1 を取るベルヌーイ確率変数である。これは説明変数に依存していないので γ の実際の値は $\mathbf{0}$ である事に注意しよう。初期値を

```
G0=matrix(0,K,1);
```

によって発生させよう。

```
Ghat=panelBLMLE(G0,Y,10)
```

とすると、γ の推定値が出力されるはずである。ここでは

```
> Ghat
          [,1]
[1,] -0.1013313
[2,] -0.1054363
[3,] -0.1445365
```

のような推定値を得られた(真の値は全て 0)。次に $\tilde{\gamma}$ を求めるプログラムを書いてみよう。(少し長いが)例えば、以下のようなプログラムによって計算できる。

```
panelBcBLMLE=function(G0,Y,N){
 Y=as.matrix(Y)
 T=nrow(Y)/N; K=ncol(Y)-1;
 Ghat=matrix(panelBLMLE(G0,Y,N),ncol=1);

 YT=matrix(0,nrow=T,ncol=N*(K+1));
 GT=matrix(0,nrow=K,ncol=T);

 for(i in 1:N){
   YT[1:T,(1+(K+1)*(i-1)):((K+1)*i)]=Y[(1+T*(i-1)):(T*i),];
  }
 for(s in 1:T){
   if (s==1){ YS=YT[2:T,]
   }else if(s==T){ YS=YT[1:T-1,]
```

```
    }else{ YS=rbind(YT[1:(s-1),],YT[(s+1):T,]); }

  YS=matrix(YS,nrow=T-1,ncol=(N*(K+1)));
  Ys=matrix(0,nrow=N*(T-1),ncol=K+1);
  for (i in 1:N){
   Ys[(1+(T-1)*(i-1)):((T-1)*i),]=YS[,(1+(K+1)*(i-1)):((K+1)*i)];
  }
  Gshat=panelBLMLE(G0,Ys,N);
  GT[,s]=matrix(Gshat,nrow=K,ncol=1);
  }

  iT=matrix(rep(1,T),nrow=T);
  Gtil=T*Ghat-((T-1)/T)*(GT%*%iT);
  return(Gtil)
 }
```

のように書ける。先ほどと同様に、これを用いて先ほどの Y のデータに実際に適用すると

```
 > Gtil=panelBcBLMLE(G0,Y,10)
 > Gtil
            [,1]
 [1,] -0.006154568
 [2,] -0.069675937
 [3,] -0.128419241
```

となった。

　また $\tilde{\gamma}$ を計算したら、それを用いて δ を推定するプログラムも書ける。この推定量を $\tilde{\delta}$ としよう。これは先ほどの対数尤度関数において γ が与えられているとして δ だけの関数として尤度関数を書き直してあげてそれを最大化するようにすればよい。対数尤度を計算する関数を panelBLFE それを固定効果パラメーターについて最大化する関数を panelBLmleFE とすると、例えば以下のようなコードが考えられる。

```
panelBLFE=function(Y,Gam,N){
 T=nrow(Y)/N; K=ncol(Y)-1;

 X=matrix(Y[,2:(K+1)],nrow=N*T);
```

```
 y=matrix(Y[,1],nrow=N*T);
 iT = matrix(rep(1,T),nrow=T);

 function(Del){
  Z=Del%x%iT + X%*%Gam;
  f=sum(y*Z - log(1+exp(Z)));
  y=-f; return(y)
 }
}
```

および

```
panelBLmleFE=function(Y,Gam,N){
 Y=as.matrix(Y);
 Del0= matrix(rep(0,N),ncol=1);
 Deltil=optim(Del0,panelBLFE(Y,Gam,N),gr=NULL,method=c("BFGS"),
 control=list(maxit=100000,ndeps=(1e-5*rep(1,N)),reltol=1e-10));
 Deltil=matrix(Deltil$par[1:N],N,1);
 return(Deltil);
}
```

これをγとして、先ほど推定値 $\tilde{\gamma}$ を用いて実際に走らせてみると

```
> Deltil=panelBLmleFE(Y,Gtil,10)
> Deltil
            [,1]
 [1,] -16.0971274
 [2,]   0.4011151
 [3,]  -0.3741035
 [4,]  -0.5405598
 [5,]  -1.5486720
 [6,]   0.4066457
 [7,]   0.3309109
 [8,]  -0.3607678
 [9,]   1.3333341
[10,]   1.3988510
```

となった。

以上のようにパネル二項ロジットモデルで、バイアス修正した推定量を推定する R のプログラムを書いたが、これらは尤度関数からその最大化を行うコードまで全部書いたものである。また推定においては尤度関数のパラメーターについての数値微分を用いる方法を採用している。

固定効果モデルはダミー変数をいれれば通常の 2 項ロジットモデルの推定に用いた R の package の関数がそのまま使えるので、これを利用してプログラムを書いてもよい。そのようなプログラムの例として mlogit 関数を用いたものをゼミのページに置いておく。mlogit 関数は数値微分ではなく解析的な微分を用いて推定しているので、先ほどのプログラムよりも若干正確だろうと思われる。ただしこのプログラムは元のデータが式(1)のようなかたちではなく、データ Y は

```
> head(Y)
  ID y        X1          X2         X3
1  1 1 -0.2332317 -0.45146442 -0.4446888
2  1 0 -0.2265353  0.79096560  0.6123357
3  1 0 -0.9964734  0.51921276 -0.7079349
4  1 1 -0.9247131 -0.01682171  0.4477253
5  1 1  1.8992607 -0.96538213  0.5599539
6  2 0 -0.7485809  1.09134378 -0.4759389
```

のように、最初の列に個人の番号が入ったようなデータフレームの形で与えられている事を想定している。1 行目が変数の名前である。ここでは ID が個人の番号、y が被説明変数、X1, X2, X3 が 3 つの説明変数の名前である。これらの名前は別に何でもよいが、1 列目に個人の番号をいれる必要がある。以下に使い方を説明する。

前提として mlogit パッケージはインスツールしてあり読み込んであるとする。まずゼミページの panelMLest.R というファイルにある "## Functions for estimation" という項目の関数を全て読み込む。バイアスを修正しない推定は(固定効果のダミーを入れて推定しただけのもの)"panelMLmle"という関数で行う事ができる、以下のように入力する

```
panelMLmle(Y,"A","B","C","D","E")
```

ここで Y はデータ名、A はデータ内の被説明変数の名前、B,C,D は推定において mlogit 関数を用いているのでそれに準じた入力で(実は上記の関数はパネルデータにおいて多項ロジットモデルにダミーをいれて推定するようになっている)、B は係数は選択肢に依存せず説明変数が選

択肢と個人の両方に依存する説明変数の名前(を+で結んだもの)、C は係数が選択肢に依存するが説明変数は個人だけに依存している説明変数の名前(を+で結んだもの)、D は係数は選択肢に依存し、説明変数は選択肢と個人の両方に依存する説明変数の名前(を+で結んだもの)、E は mlogit 関数の reflevel を指定する選択肢、となっている。これらの説明について、よくわからない場合は mlogit 関数のゼミ資料をもう一度読むとよい。

例えば、先ほどのようにデータとして Y が与えられており、説明変数として X1 と X2 だけ使用する場合、推定値に Ghat と名前を付けるとして、

```
> C="X1+X2"
> Ghat=panelMLmle(Y,"y","0",C,"0","0")
```

と入力すればよい(上記の最後の引数(reflevel)は 0 とした)。ここで mlogit 関数と同様、対応する変数がない場合は変数名のところを 0 とする。この関数は mlogit 関数を用いて固定効果のダミー変数をいれたものを推定している。

```
> summary(Ghat)
```

とすれば推定結果を見ることができる(固定効果の推定値は i 番目の個人に対して fi の係数として出力される)。また panelBcMLmle はジャックナイフ法によってバイアスを修正したもの(ただし推定は mlogit 関数で行ったもの)であり、panelMLmle と同様に使用できる。推定値に Gtil と名前を付け、先ほどのデータに対して推定すると

```
> Gtil=panelBcMLmle(Y,"y","0",C,"0","0")
```

のように入力すればよい(以下の固定効果推定のコードと整合的にするには、上記の最後の引き数は"0"にしないといけない)。またこの時、固定効果を推定するには panelMLmleFE を用いて

```
> Deltil=panelMLmleFE(Y,Gtil,C)
```

とする。

参考文献

Hahn and Newey (2004) “Jackknife and Analytical Bias Reduction for Nonlinear Panel Models,” *Econometrica*, 72(4), pp.1295-1319.