代理変数モデルの分析

平　館　道　子

代理変数モデルの分析

経済理論を実証しようとするとき，しばしば，理論上の変数に関するデータが利用不可能であったり，観測不可能であったりするという事態に遭遇する。このような場合その変数を無視できないと感じるならば，私達はその代理をするであろうと思われる変数を用いる。例えば回帰モデルの場合，観測不可能な説明変数に関して代理変数を用いる等である。理論の実証は結局観測可能な現象においてなさざるを得ないから，代理変数として用いられるものは観測不可能な仮説的変数に少くとも確率的に関連づけられていなければならないが，その関連を現実に確かめることはできない。したがってデータの証拠を用いてそれを推定することができるかが1つの問題である。もう1つの問題は無論パラメータに関する推測である。代理変数による回帰式の係数は，真の変数による回帰式の係数とは一致しないであろうが，もし代理変数が真の変数と高い相関をもつならば，差は小さいであろう。回帰係数の符号を変化させないことを保証するほど相関が高ければ，正しい推測が行われるだろうが，それではどの程度相関は高くなければならないか。この問題に対して，Krasker-Pratt(1986)は代理変数が1つの場合について回答を与え，さらに一般の場合のアルゴリズムを提示している。

実際の分析では，代理変数としていくつかの変数をとり上げ，結果を得てそれが満足の行くものでない時には，もっと良い代理変数を見つけることができればもっと満足の行く結果が得られるであろう，と考えるであろう。このような場合理論の実証とはどのような事になるのであろうか。良い代理変

数を求めつづけるならば理論は反証されることはないのである。

1．単純な代理変数モデル

第1のモデルは次の様なものである。

$$Y_i=\beta\chi_i+u_i \tag{1}$$

$$x_i=\delta\chi_i+\varepsilon_i. \tag{2}$$

χ_i は観測不可能な変数であり，式（1）は従属変数 Y_i が χ_i を説明変数とする一次関数と誤差項 u_i とから成ることを示している。式（2）は x_i が χ_i の代理変数であることを示し，$\delta>0$ である。もし $\delta=1$ ならば x_i は χ_i の測定値であり，測定誤差 ε_i を伴う〝変数の誤差〟モデルとなる。したがって，代理変数問題と変数の誤差問題とは共通の性格をもっている。u_i, ε_i はそれぞれ独立に平均0，分散 σ_u^2, σ_ε^2 の正規分布に従うと想定する。理論的に興味がもたれるパラメータは β である。

このモデルは識別性に欠けている。なぜならば，パラメータ $(\beta,\ \delta)$ の尺度変化が χ の尺度変化によって相殺されるために，観測可能な変数 $(Y_i,\ x_i)$ の分布は $(\beta,\ \delta)$ の尺度変化によって影響を受けないからである。すなわち，構造は異っても観測値の分布は同じである。したがって $(\beta,\ \delta)$ は尺度要因を除いて一意的である。Y の x に関する回帰は，δ の符号が正であることが与えられる時，β の符号のみを決定できる。

2．特異性

$\mathbf{Y}'=(Y_1,Y_2,\cdots\cdots,Y_N)$, $\mathbf{x}'=(x_1,x_2,\cdots\cdots,x_N)$, $\boldsymbol{\chi}'=(\chi_1,\chi_2,\cdots\cdots,\chi_N)$ とすると尤度関数は次の様になる。

$$\begin{aligned}L(\boldsymbol{\chi},\beta,\delta,\sigma_u^2,\sigma_\varepsilon^2|\mathbf{Y},\mathbf{x})&\propto(\sigma_u^2)^{-\frac{N}{2}}\exp\left[-\frac{1}{2\sigma_u^2}(\mathbf{Y}-\boldsymbol{\chi}\beta)'(\mathbf{Y}-\boldsymbol{\chi}\beta)\right]\\&\times(\sigma_\varepsilon^2)^{-\frac{N}{2}}\exp\left[-\frac{1}{2\sigma_\varepsilon^2}(\mathbf{x}-\boldsymbol{\chi}\delta)'(\mathbf{x}-\boldsymbol{\chi}\delta)\right]\end{aligned} \tag{3}$$

この尤度関数は次の様な性格をもつ。いま尤度関数の等高線を考えることにして，対数尤度を定数cに等しいとおくと次の様になる。

$$-N\sigma_u^2\sigma_\varepsilon^2\log\sigma_u^2-\sigma_\varepsilon^2(\mathbf{Y}-\boldsymbol{\chi}\beta)'(\mathbf{Y}-\boldsymbol{\chi}\beta)$$
$$-N\sigma_u^2\sigma_\varepsilon^2\log\sigma_\varepsilon^2-\sigma_u^2(\mathbf{x}-\boldsymbol{\chi}\delta)'(\mathbf{x}-\boldsymbol{\chi}\delta)=4\sigma_\varepsilon^2\sigma_u^2c.$$
$$\sigma_\varepsilon^2(\mathbf{Y}-\boldsymbol{\chi}\beta)'(\mathbf{Y}-\boldsymbol{\chi}\beta)+\sigma_u^2(\mathbf{x}-\boldsymbol{\chi}\delta)'(\mathbf{x}-\boldsymbol{\chi}\delta)$$

$$=-N\sigma_u^2\sigma_\varepsilon^2\log\sigma_u^2-N\sigma_u^2\sigma_\varepsilon^2\log\sigma_\varepsilon^2-4\sigma_\varepsilon^2\sigma_u^2c \tag{4}$$

σ_u^2 を0にすると（4）の右辺はcのどのような値に対しても0になる。したがって

$$\sigma_\varepsilon^2(Y-\chi\beta)'(Y-\chi\beta)=0$$

となり，$\sigma_\varepsilon^2\neq 0$ であれば $\chi=\frac{Y}{\beta}$ である。つまりすべての等高線は $\sigma_u^2=0$, $\chi=\frac{Y}{\beta}$ という点を通ることになる。また $\sigma_\varepsilon^2=0$ とすると同様にcの値にかかわらず $\chi=\frac{x}{\delta}$ となり，すべての等高線はこの点をも通ることになる。したがってこの2つの点の近くでは尤度関数は任意の正の値をとり得るのである。この事を，尤度関数は $(0,\ \sigma_\varepsilon^2,\ \frac{Y}{\beta},\ \beta,\ \delta)$ と $(\sigma_u^2,\ 0,\ \frac{x}{\delta},\ \beta,\ \delta)$ において本質的な特異性をもつ，という。尤度関数はこれらの点の上ではよく定義されないのである。いまこれら2組の制約条件の下に尤度関数を最大化すると，それぞれ，$\widehat{\left(\frac{\beta}{\delta}\right)}=(x'x)^{-1}x'Y$ と，$\widehat{\left(\frac{\beta}{\delta}\right)}=(Y'x)^{-1}Y'Y=((Y'Y)^{-1}Y'x)^{-1}$ という推定値が得られる。前者はYのxに関する回帰であり，後者はxのYに関する回帰の逆数である。

3．構造形モデル

次に，観測不可能な χ に対して分布が想定されるとしよう。これは構造形といわれるモデルである。χ の分布がモデルに属するものか，事前分布であるのかは，ベイジアンにとってはあまり問題ではない。構造モデルと呼ばれているものでは通常 χ の分布を，平均 $\bar{\chi}$，分散 σ_χ^2 の正規分布であると想定するが，Zellner (1971) ではもっと他の事前分布についても分析されている。通常の想定にしたがって $\chi\sim N(\bar{\chi},\ \sigma_\chi^2)$ とすると，$(Y_i,\ x_i)$ は平均が $(\beta\bar{\chi},\ \delta\bar{\chi})$，分散 Σ が次の様な2変量正規分布に従う。

$$\Sigma=\begin{pmatrix}\beta^2\sigma_\chi^2+\sigma_u^2 & \delta\beta\sigma_\chi^2\\ \delta\beta\sigma_\chi^2 & \delta^2\sigma_\chi^2+\sigma_\varepsilon^2\end{pmatrix}.$$

この事から，x_i が与えられた時の Y_i の条件つき分布は平均が $\frac{\delta\beta\sigma_\chi^2}{\delta^2\sigma_\chi^2+\sigma_\varepsilon^2}x_i$ の正規分布であることがわかる。すなわち，Yをxに関して回帰すると，

$\dfrac{\beta}{\delta}$ の縮少された係数 $\dfrac{\beta}{\delta}\left[\dfrac{\sigma_\chi^2}{\sigma_\chi^2+\sigma_\varepsilon^2/\delta^2}\right]$ が得られることが示唆される。また同様にxのYが所与の時の平均は $\dfrac{\delta\sigma_u^2}{\beta\sigma_\chi^2+\sigma_u^2/\beta}\mathrm{Y}$ であるから，xをYに関して回帰すると $\dfrac{\delta}{\beta}$ の縮小された係数 $\dfrac{\delta}{\beta}\left[\dfrac{\sigma_\chi^2}{\sigma_\chi^2+\sigma_u^2/\beta^2}\right]$ が得られる。前者では σ_ε^2 が $\delta^2\sigma_\chi^2$ に比して小さければ縮小のファクターは1に近くなり，後者では σ_u^2 が $\beta^2\sigma_\chi^2$ に比して小さければ1に近づく。この様に回帰係数が $\dfrac{\beta}{\delta}$ のみの関数になっていることは識別性欠除の結果に他ならない。

構造形の尤度関数は変数が平均のまわりで定義されているとして，

$$S=\begin{bmatrix}\mathbf{Y'Y} & \mathbf{Y'x}\\ \mathbf{x'Y} & \mathbf{x'x}\end{bmatrix}$$

とすると次の様になる。

$$\mathrm{L}(\beta,\delta,\sigma_u^2,\sigma_\varepsilon^2,\sigma_\chi^2|\mathbf{Y},\mathbf{x})\propto|\Sigma|^{-\frac{N}{2}}\exp\left\{-\frac{1}{2}\sum_i(\mathrm{Y}_i,\mathrm{x}_i)'\Sigma^{-1}(\mathrm{Y}_i,\mathrm{x}_i)\right\}$$

$$=\left|\Sigma\right|^{-\frac{N}{2}}\exp\left\{-\frac{1}{2}\mathrm{tr}\Sigma^{-1}\mathrm{S}\right\}. \qquad (5)$$

$\dfrac{\beta}{\delta}$ の順回帰による推定値 $\left(\widehat{\dfrac{\beta}{\delta}}\right)^{\mathrm{D}}$ を $\hat{\theta}^{\mathrm{D}}$, 逆回帰によるそれ $\left(\widehat{\dfrac{\beta}{\delta}}\right)^{\mathrm{R}}$ を $\hat{\theta}^{\mathrm{R}}$ とすると，尤度関数について次の様なことがいえる。

尤度関数（5）は $\hat{\theta}^{\mathrm{D}}$ と $\hat{\theta}^{\mathrm{R}}$ の間の任意の $\dfrac{\beta}{\delta}$ の値において最大値を達成する。

証明：Σ に制約がない場合，この尤度関数が $\hat{\Sigma}=\dfrac{\mathrm{S}}{\mathrm{N}}$ において最大値 $\left|\dfrac{\mathrm{S}}{\mathrm{N}}\right|^{-\frac{N}{2}}\exp\{-\mathrm{N}\}$ をとることはよく知られている。このことから $\hat{\Sigma}=\dfrac{\mathrm{S}}{\mathrm{N}}$ と矛盾しない β, δ, σ_u^2, σ_ε^2, σ_χ^2 の値が存在すれば，Σ のこの値は実行可能であり，そのような $(\beta, \delta, \sigma_u^2, \sigma_\varepsilon^2, \sigma_\chi^2)$ の値は最尤推定値である。Σ と $\hat{\Sigma}$ とを対比することによって次の各式が得られる。

$$\beta\delta\sigma_\chi^2=\frac{\mathbf{x'Y}}{\mathrm{N}},$$

$$\beta^2\sigma_\chi{}^2+\sigma_u{}^2=\frac{\mathbf{Y'Y}}{N},$$

$$\delta^2\sigma_\chi{}^2+\sigma_\varepsilon{}^2=\frac{\mathbf{x'x}}{N}.$$

これから,

$$\sigma_\chi{}^2=\frac{\mathbf{x'Y}}{\beta\delta N},$$

$$\sigma_u{}^2=\frac{\mathbf{Y'Y}}{N}-\beta^2\sigma_\chi{}^2=\frac{1}{N}\left[\mathbf{Y'Y}-\left(\frac{\beta}{\delta}\right)\mathbf{x'Y}\right],$$

$$\sigma_\varepsilon{}^2=\frac{\mathbf{x'x}}{N}-\delta^2\sigma_\chi{}^2=\frac{1}{N}\left[\mathbf{x'x}-\left(\frac{\delta}{\beta}\right)\mathbf{x'Y}\right]$$

となる。$\sigma_u{}^2$, $\sigma_\varepsilon{}^2\geq 0$ であることから,

$$\mathbf{Y'Y}-\frac{\beta}{\delta}\mathbf{x'Y}\geq 0,$$

$\mathbf{x'Y}\gtreqless 0$ のとき　$\left(\frac{\beta}{\delta}\right)\lesseqgtr(\mathbf{x'Y})^{-1}\mathbf{Y'Y}=\hat{\theta}^R.$

$$\mathbf{x'x}-\frac{\delta}{\beta}\mathbf{x'Y}\geq 0,$$

同様に $\mathbf{x'Y}$ の符号に従って,

$$\left(\frac{\beta}{\delta}\right)^{-1}\lesseqgtr(\mathbf{x'Y})^{-1}(\mathbf{x'x}).\ \text{したがって}\ \left(\frac{\beta}{\delta}\right)\gtreqless(\mathbf{x'x})^{-1}\mathbf{x'Y}=\hat{\theta}^D$$

が導びかれる。すなわち, $\hat{\theta}^D$ と $\hat{\theta}^R$ の間の $\frac{\beta}{\delta}$ の値に対して尤度関数は最大値をとる。

4. 他の説明変数がある場合

次に

$$Y_i=\beta\chi_i+\gamma z_i+u_i \tag{6}$$

$$x_i=\delta\chi_i+\varepsilon_i \tag{7}$$

というモデルを考察しよう。z_i はもう1つの説明変数であり観測可能である。前の場合と同様に特異性は $\sigma_u{}^2=0$, 又は $\sigma_\varepsilon{}^2=0$ によって生じる。前者の場合 $Y_i=\beta\chi_i+\gamma z_i$ であるから, $x_i=\frac{\delta}{\beta}(Y_i-\gamma z_i)+\varepsilon_i$ であり, 後者の場合 $x_i=\delta\chi_i$ であるから, $Y_i=\frac{\beta}{\delta}x_i+\gamma z_i+u_i$ である。これらのことから γ の推定

は可能であることがわかる。前者の場合xをYとzに関して回帰させ，それをYについて解き直すことでγを推定することができるであろう。また後者の場合はYをxとzに回帰させることによってγを推定できるであろう。前者は逆回帰による推定であり，これを$\hat{\gamma}^{R}$としよう。後者は順回帰であり，これを$\hat{\gamma}^{D}$としよう。ここに

$$\hat{\gamma}^{D}=\frac{x'xz'Y-x'zx'Y}{x'xz'z-(x'z)^2},\quad \hat{\gamma}^{R}=\frac{Y'zx'Y-Y'Yz'x}{z'zx'Y-z'Yz'x}$$

である。

構造モデルはχの分布を想定しなければならないので，χ_iは$N(\kappa z_i,\ \sigma_\chi^2)$という正規分布に従うとしよう。$\chi$とzが相関をもたなければ，Yをzに回帰させて単純回帰によってγを推定することができ，χの代理変数を探す動機はなくなる。χとzが相関をもてば単純回帰は，zと相関のある変数を除外したことによって，偏った推定値を与えることになるから，χの代理変数を求めざるを得ないのである。

z_iを与えた時の$(Y_i,\ x_i)$の条件つき分布は2変量正規分布であり，その平均と分散は

$$E[(Y_i,x_i)|z_i]\ =\ [(\beta\kappa+\gamma)z_i,\delta\kappa z_i],$$

$$V[(Y_i,x_i)|z_i]\ =\ \begin{bmatrix}\beta^2\sigma_\chi^2+\sigma_u^2 & \beta\delta\sigma_\chi^2\\ \beta\delta\sigma_\chi^2 & \delta^2\sigma_\chi^2+\sigma_\varepsilon^2\end{bmatrix}\equiv\Sigma$$

となる。これが条件つき分布であることに注意して，

$$S\ =\ \begin{pmatrix}\dfrac{Y'Yz'z-(z'Y)^2}{z'z} & \dfrac{x'Yz'z-z'Yz'x}{z'z}\\[2ex] \dfrac{x'Yz'z-z'Yz'x}{z'z} & \dfrac{x'xz'z-(z'x)^2}{z'z}\end{pmatrix}$$

$$\widehat{EY}_i=\frac{z'Y}{z'z}z_i,\quad \widehat{Ex}_i=\frac{z'x}{z'z}z_i$$

とすると，前節と同様な議論によって，尤度関数は$\hat{\Sigma}=\frac{S}{N}$，$\widehat{E(Y_i|z_i)}=\widehat{EY}_i$，$\widehat{E(x_i|z_i)}=\widehat{Ex}_i$の時最大値をとり，ここから$\frac{\beta}{\delta}$が$\hat{\theta}^{D}$と$\hat{\theta}^{R}$の間で尤度が最大になることが導びかれる。更にこの時期待値の関係に注意すると，

$$\hat{\gamma}^{D}=\frac{\mathbf{Y'z}}{\mathbf{z'z}}-\frac{\mathbf{x'z}}{\mathbf{z'z}}\hat{\theta}^{D}$$

$$\hat{\gamma}^{R}=\frac{\mathbf{Y'z}}{\mathbf{z'z}}-\frac{\mathbf{x'z}}{\mathbf{z'z}}\hat{\theta}^{R}$$

から，γ が $\hat{\gamma}^{D}$ と $\hat{\gamma}^{R}$ の間の任意の値をとるとき尤度が最大になることが導びかれる。更に $\hat{\gamma}^{S}$ を Y を z に回帰させた時の γ の単純推定値とすると，

$\hat{\gamma}^{S}=\dfrac{\mathbf{z'Y}}{\mathbf{z'z}}$ であるから，

$$\hat{\gamma}^{D}=\hat{\gamma}^{S}-\frac{\mathbf{x'z}}{\mathbf{z'z}}\hat{\theta}^{D}=\hat{\gamma}^{S}-\frac{\mathbf{x'z}}{\mathbf{z'z}}\hat{\theta}^{R}R^{2}_{xY\cdot z}$$

$$\hat{\gamma}^{R}=\hat{\gamma}^{S}-\frac{\mathbf{x'z}}{\mathbf{z'z}}\hat{\theta}^{R}$$

となるので，（$R^{2}_{Yx\cdot z}$ はYをxとzに回帰させた時の偏相関係数の２乗である）

$$\hat{\gamma}^{D}-\hat{\gamma}^{S}=R^{2}_{xY\cdot z}(\hat{\gamma}^{R}-\hat{\gamma}^{S})$$

が成立つ。したがって，$\hat{\gamma}^{D}$ は $\hat{\gamma}^{S}$ と $\hat{\gamma}^{R}$ の間にあり，$R^{2}_{xY\cdot z}$ が小さければ $\hat{\gamma}^{D}$ は $\hat{\gamma}^{S}$ から十分には離れていないことがわかり，誤差を含む変数を説明変数として用いると，誤差を含まない変数の係数の推定値は，誤差を含む変数を無視して省略してしまう場合にくらべて，それほど大きな差がないことがわかる。

5．操作変数モデル

次に代理変数の他に測定誤差を伴う変数が存在する場合を考えよう。モデルは次の様である。

$$Y_i=\beta\chi_i+u_i \tag{8}$$

$$x_i=\delta\chi_i+\varepsilon_{1i} \tag{9}$$

$$w_i=\chi_i+\varepsilon_{2i} \tag{10}$$

w_i は χ_i の測定値であり，x_i は χ_i の代理変数である。u，ε_1，ε_2 は独立に平均０，分散が σ_u^2，σ_1^2，σ_2^2 の正規分布に従う確率変数であるとする。もし代理変数xに誤差がなければ（$\sigma_1^2=0$），$\chi_i=\dfrac{x_i}{\delta}$ であるから，

$$Y_i=\frac{\beta}{\delta}x_i+u_i$$

$$w_i=\frac{x_i}{\delta}+\varepsilon_{2i}$$

となる。したがってYをxに回帰することによって $\frac{\beta}{\delta}$ を $(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{Y}$ によって推定し，wをxに回帰することによって $\frac{1}{\delta}$ を $(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{w}$ によって推定することができ，両者の比によって β を推定することができよう。すなわち

$$\hat{\beta}^1=\frac{\mathbf{x}'\mathbf{Y}}{\mathbf{x}'\mathbf{w}}.$$

これは操作変数による推定値であり，wは操作変数と呼ばれているものである。すなわち操作変数モデルは変数の誤差モデルと代理変数モデルのあいの子なのである。

6．複数の代理変数のあるモデル

観測できない変数 χ に対して代理変数が複数ある場合について考察しよう。この場合推定値の領域を狭めることができるかどうかが問題である。構造モデルは次の様である。

$$Y_i=\beta\chi_i+\gamma z_i+u_i \tag{11}$$

$$\mathbf{x}_i=\boldsymbol{\delta}\chi_i+\boldsymbol{\varepsilon}_i \tag{12}$$

$$\chi_i=\varkappa z_i+e_i \tag{13}$$

$\mathbf{x}_i$ はp次元の代理変数のベクトルであり，$\boldsymbol{\delta}$ は同様にp次元ベクトルである。u_i, e_i は平均0，分散がそれぞれ σ_u^2, σ_χ^2 の正規分布に独立に従っていると想定する。$\boldsymbol{\varepsilon}$ は平均 $\mathbf{0}$，共分散行列 Ω の p 次元正規分布に従う。Ω は $p\times p$ 行列である。(11) を用いて χ について積分すると，観測可能な変数 $(Y_i, \mathbf{x}_i')$ の z_i が与えられた時の条件つき分布は平均，共分散が次の様な正規分布である。

$$E[(Y_i,\mathbf{x}_i')|z_i]=(\beta\varkappa z_i+\gamma z_i, \varkappa z_i\boldsymbol{\delta}') \tag{14}$$

$$V[(Y_i,\mathbf{x}_i')|z_i]=\begin{bmatrix}\beta^2\sigma_\chi^2+\sigma_u^2 & \beta\sigma_\chi^2\boldsymbol{\delta}' \\ \boldsymbol{\delta}\beta\sigma_\chi^2 & \sigma_\chi^2\boldsymbol{\delta}\boldsymbol{\delta}'+\Omega\end{bmatrix}.$$

この場合も識別問題を避けることはできない。何故なら $(\beta, \boldsymbol{\delta}')$ の変換は $(\varkappa, \sigma_\chi^2)$ の変化によって相殺され，この分布は少しも変化しないからである。

$M=I-\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'$, X はp個の代理変数の観測値の行列 $(N\times p)$ とすると，

この分布が条件つき分布であることに注意すれば前と同様に，

$$S=\begin{bmatrix} Y'MY & Y'MX \\ X'MY & X'MX \end{bmatrix}$$

$$\widehat{EY}_i=(z'z)^{-1}z'Yz_i$$

$$\widehat{EX}_i=(z'z)^{-1}z'xz_i$$

において

$$\hat{\Omega}=\frac{S}{N}$$

のとき，尤度関数は最大値をとるから，最尤推定値は次の様な方程式を満足する。

$$\beta x+\gamma=(z'z)^{-1}z'X$$

$$\beta^2\sigma_x{}^2+\sigma_u{}^2=\frac{Y'MY}{N}$$

$$\delta\sigma_x{}^2\delta'+\Omega=\frac{X'MX}{N}$$

$$\beta\sigma_x{}^2\delta'=\frac{Y'MX}{N}. \tag{15}$$

次にある1つの x，x_l をYとzに回帰させ，$\dfrac{\beta}{\delta_l}$ の逆回帰による推定値を求めると，

$$\widehat{\left(\frac{\beta}{\delta_l}\right)}^R=\frac{Y'MY}{Y'Mx_l}$$

となる。また γ に関する同様の推定値 $\hat{\gamma}^R$ は同様にして次の様になることが得られる。

$$\hat{\gamma}_l{}^R=(z'z)^{-1}z'Y-(z'z)^{-1}z'x_l\widehat{\left(\frac{\beta}{\delta_l}\right)}^R. \tag{16}$$

そこで $(Y_i'x_i')$ のzが与えられた時の条件つき分布においては，回帰係数 $xz_i\delta'$ と共分散 $\beta\sigma_x{}^2\delta'$ が δ に比例していることに注意して，もし標本積率にもこの関係があると仮定するならば，すなわち

$$(z'z)^{-1}z'X\propto Y'MX \tag{17}$$

と仮定するならば，γ のすべての逆回帰推定値は等しいことがわかる。それは以下の通りである。

比例条件を個々の x_l と x_j について書くと，

$$\frac{(z'z)^{-1}z'x_l}{Y'Mx_l}=\frac{(z'z)^{-1}z'x_j}{Y'Mx_j}$$

となるので，

$$z'x_l\left(\frac{\hat{\beta}}{\delta_l}\right)^{R}=z'x_j\left(\frac{\hat{\beta}}{\delta_j}\right)^{R}$$

が成立つ．したがって (16) より

$$\hat{\gamma}_l^{R}=\hat{\gamma}_j^{R}\ (\equiv\hat{\gamma}^{R})$$

が成立つ。さらにこの関係を用いれば次の結果が得られる。

もし比例条件 (17) が成立っているならば尤度関数を最大にする γ の値は，$\hat{\gamma}^{R}$ と $\hat{\gamma}^{D}$ の間にある。ここに $\hat{\gamma}^{D}$ は Y を z とすべての代理変数 x に回帰させた時の γ の推定値である。

証明：任意の β と δ_1 の値をえらぶと，まず $\hat{\gamma}^{D}$ は $(z'z)^{-1}z'Y-(z'z)^{-1}z'x_1\left(\widehat{\frac{\beta}{\delta_1}}\right)^{R}R^2_{Yx\cdot z}$ である。次に方程式 (15) を他のパラメーターについて解くと，

$$\chi=\frac{(z'z)^{-1}z'x_1}{\delta_1},$$

$$\sigma_\chi{}^2=\frac{Y'Mx_1}{\beta\delta_1 N},$$

$$\sigma_u{}^2=\frac{Y'MY}{N}-\frac{\beta Y'Mx_1}{\delta_1 N},$$

$$\gamma=(z'z)^{-1}z'Y-\frac{\beta}{\delta_1}(z'z)^{-1}z'x_1,$$

$$\begin{aligned}\Omega&=\frac{X'MX}{N}-\delta\sigma_\chi{}^2\delta'\\&=\frac{X'MX}{N}-\frac{X'MYY'MX}{N^2\beta^2\sigma_\chi{}^2}\quad\left(\because\quad \delta'=\frac{Y'MX}{N\beta\sigma_\chi{}^2}\right)\\&=\frac{X'MX}{N}-\frac{\delta_1}{Y'Mx_1\beta N}X'MYY'MX.\end{aligned}$$

$\sigma_u{}^2>0$ であるから

$$\frac{\beta}{\delta_1}<\frac{Y'MY}{Y'Mx_1}=\left(\widehat{\frac{\beta}{\delta_1}}\right)^{R}.$$

Ω が正値定符号であることから，

$$\frac{\delta_1}{Y'Mx_1\beta} < (X'MY)^{-1}(X'MX)(Y'MX)^{-1}$$
$$= [Y'MX(X'MX)^{-1}(X'MY)]^{-1}$$

したがって

$$\frac{\beta}{\delta_1} > \frac{Y'MX(X'MX)^{-1}X'MY}{Y'Mx_1}$$
$$= \widehat{\left(\frac{\beta}{\delta_1}\right)}^{R} \cdot R^2.$$

ここに R^2 は z が与えられる時に Y をすべての代理変数について回帰した時の多重相関係数の 2 乗 $R^2_{Yx\cdot z}$ である。したがって,

$$(z'z)^{-1}z'Y - (z'z)^{-1}z'x_1\widehat{\left(\frac{\beta}{\delta_1}\right)}^{R}R^2 > \hat{\gamma} > (z'z)^{-1}z'Y - (z'z)^{-1}z'x_1\widehat{\left(\frac{\beta}{\delta_1}\right)}^{R}$$

が導かれ，この不等式の左辺が $\hat{\gamma}^D$ であることに注意すれば,

$$\hat{\gamma}^D > \hat{\gamma} > \hat{\gamma}^R$$

が導びかれる。

更に $\hat{\gamma}^S = (z'z)^{-1}z'Y$ とすると前と同様に

$$(\hat{\gamma}^D - \hat{\gamma}^S) = R^2(\hat{\gamma}^R - \hat{\gamma}^S)$$

が導びかれる。

この結果から，代理変数が 1 つの場合にくらべて複数個ある場合には，多重相関係数 R は偏相関係数 $R_{xY\cdot z}$ より大きくなるから，$\hat{\gamma}$ の限界はより狭くなることがわかる。

7．変数の誤差モデルにおける事後分布

代理変数問題は変数の誤差問題と深くかかわっていることは既に述べた通りである。変数の誤差あるいは測定誤差モデルは次の様なものである。

$$Y_i = \beta\chi_i + u_i \tag{18}$$

$$x_i = \chi_i + \varepsilon_i \tag{19}$$

u_i, ε_i はこれまでと同様平均 0，分散 σ_u^2, σ_ε^2 の正規分布に独立に従うと想定される。さらに χ が平均 $\bar{\chi}$, 分散 σ_χ^2 の正規分布に従うと想定するとき構造モデルになる。このモデルについてベイジアンからの接近法を用いて，Lindley-El-Sayyad (1968) が近似的な結果を求めているのでその概略について述べておこう。このモデルの構造形においては，(Y_i, x_i) の分布は平均が

$(\beta\chi, \ \chi)$ で共分散行列 Σ が

$$\Sigma=\begin{bmatrix} \sigma_u^2+\beta^2\sigma_\chi^2 & \beta\sigma_\chi^2 \\ \beta\sigma_\chi^2 & \sigma_\chi^2+\sigma_\epsilon^2 \end{bmatrix}$$

の２変量正規分布である。ここで $\bar{\chi}=0$ と仮定し，$(\beta, \ \sigma_u^2, \ \sigma_\epsilon^2, \ \sigma_\chi^2)$ の事前分布を $f(\beta, \ \sigma_u^2, \ \sigma_\epsilon^2, \ \sigma_\chi^2)$ とすると事後分布は次の様になる。

$$f(\beta,\sigma_u^2,\sigma_\epsilon^2,\sigma_\chi^2|\mathbf{Y},\mathbf{x})\propto\frac{f(\beta,\sigma_u^2,\sigma_\epsilon^2,\sigma_\chi^2)}{(\beta^2\sigma_\chi^2\sigma_\epsilon^2+\sigma_\chi^2\sigma_u^2+\sigma_\epsilon^2\sigma_u^2)^{\frac{N}{2}}}$$

$$\times\exp\left\{-\frac{N}{2}\frac{\mathbf{x}'\mathbf{x}(\beta^2\sigma_\chi^2+\sigma_u^2)-2\mathbf{x}'\mathbf{Y}\beta\sigma_\chi^2+\mathbf{Y}'\mathbf{Y}(\sigma_\chi^2+\sigma_\epsilon^2)}{\beta^2\sigma_\chi^2\sigma_\epsilon^2+\sigma_\chi^2\sigma_u^2+\sigma_\epsilon^2\sigma_u^2}\right\}.$$

ここで $\lambda=\frac{\sigma_u^2}{\sigma_\epsilon^2}$，$\rho=\frac{\sigma_\chi^2}{\sigma_\epsilon^2}$ とすると事後分布は

$$\frac{f(\beta,\sigma_\epsilon^2,\lambda,\rho)}{\sigma_\epsilon^{2N}(\lambda+\lambda\rho+\rho\beta^2)^{\frac{N}{2}}}\exp\left\{-\frac{N}{2}\frac{\mathbf{x}'\mathbf{x}(\beta^2\rho+\lambda)-2\mathbf{x}'\mathbf{Y}\beta\rho+\mathbf{Y}'\mathbf{Y}(\rho+1)}{\sigma_\epsilon^2(\lambda+\lambda\rho+\rho\beta^2)}\right\} \tag{20}$$

となる。ここで σ_ϵ^2 の事前分布を他のパラメータと独立に σ_ϵ^{-2} に比例する曖昧分布を想定して（20）を σ_ϵ^2 について積分すると，

$$f(\beta,\lambda,\rho|\mathbf{Y},\mathbf{x})\propto f(\beta,\lambda,\rho)\left[\frac{\rho(\lambda+\beta^2)+\lambda}{\{\rho(\mathbf{x}'\mathbf{x}\beta^2-2\mathbf{x}'\mathbf{Y}\beta+\mathbf{Y}'\mathbf{Y})+\lambda\mathbf{x}'\mathbf{x}+\mathbf{Y}'\mathbf{Y}\}^2}\right]^{\frac{N}{2}}$$

となる。次に ρ に関する積分を考えるが，固定された β と λ に対して，ρ の分布を，同じモードをもち，さらにモードにおいて同じ２階の対数微係数をもつ正規分布でおき替えることによって積分を近似し，さらに固定された λ に対する β の分布を近似するためにもう一度同じ方法を用いて，次の様な結果を得る。

大標本においては，λ を固定する時の β の近似的な事後分布は次の方程式（21）の根 β_λ を平均とし，分散を

$$\frac{\beta_\lambda^2(\mathbf{x}'\mathbf{x}\mathbf{Y}'\mathbf{Y}-(\mathbf{x}'\mathbf{Y})^2)}{N(\mathbf{x}'\mathbf{Y})^2}=\frac{\beta_\lambda^2}{N}\frac{1-r^2}{r^2}$$

とする正規分布である。ここに r^2 は標本相関係数である。β_λ のみたす方程式は

$$\beta^2+t\beta-\lambda=0, \ \ t=\frac{\lambda\mathbf{x}'\mathbf{x}-\mathbf{Y}'\mathbf{Y}}{\mathbf{x}'\mathbf{Y}}. \tag{21}$$

β_λ は $\mathbf{x'Y}$ と同じ符号のものである。

$\lambda=0$ のとき，$\beta_\lambda=\mathbf{Y'Y(x'Y)^{-1}}$ となり（非零の根），これは β の逆回帰から得られる推定値 $\hat{\beta}^R$ である。また λ が無限に増大する時，β_λ は $\dfrac{\mathbf{x'Y}}{\mathbf{x'x}}$ に収束する。これは直接回帰からの推定値 $\hat{\beta}^D$ である。更にこの時分散は

$$\left(\frac{\mathbf{x'Y}}{\mathbf{x'x}}\right)^2\frac{(\mathbf{x'xY'Y}-(\mathbf{x'Y})^2)}{\mathrm{N}(\mathbf{x'Y})^2}=\frac{\mathbf{x'xY'Y}-(\mathbf{x'Y})^2}{\mathrm{N}(\mathbf{x'x})^2}$$

となって，自由度を除けばこれは $\hat{\beta}^D$ の分散の通常の推定値である。更に β_λ は λ が既知である時の通常の最小2乗法と最尤法による推定値である。これは大標本においては最尤法による結果とベイジアンの方法による結果が一致するという一般にみられる結果の一例になっている。

文献

Krasker W.S. and Pratt J.W. (1986) "Bounding the Effects of Proxy Variables on Regression Coefficients," Econometrica, Vol. 54 p.p.641-655

Leamer, E.E. (1978) "Specification Searches," John Wiley & Sons

Lindley D.V. and EL-Sayyad G.M. (1968) "The Bayesian estimation of a linear functional relationship," J. Roy. Stat. Soc. B30, 190-202

Zellner, A (1971) "An Introduction to Bayesian Inference in Econometrics," John Wiley & Sons