<論説>連立方程式の種々の推定方法の基本的一体性

連立方程式の種々の推定方法の基本的一体性

今　川　正

は　じ　め　に

経済の構造方程式システム――連立方程式モデルであらわされている――のパラメーターの推定方法にはいろいろのものがある。たとえば，代表的な計量経済学のテキストであるジョンストン『計量経済学の方法』においてはつぎの方法がとりあげられている。

普通の最小2乗法，一般化最小2乗法，限定情報単1方程式法，最小分散比法，2段階最小2乗法，間接最小2乗法，完全情報最尤法，3段階最小2乗法。

しかし，ジョンストンはこれらの方法を1つの観点から統一的にながめることはしていない。ゴールドバーガー，マランボー，クリスト，タイルも同じようにそうしていない。これはチョウによって企てられた。

チョウは標準の最小2乗法を一般化することによって，上記の種々の推定方法に到達することを試みている。（それの吟味はあとでする。）われわれはチョウと逆の途をたどって，一般的な完全情報最尤

法よりはじめて，それに限定を加えてゆくことによって上記の種々の推定方法をえることを試みる。

第1節　多数方程式モデル

攪乱項ベクトルの縦配列

次節以降の展開の便宜のためにつぎの多数方程式モデルへの最尤法の適用について復習しておこう。

$$\mathbf{y}_j=\mathbf{X}_j\boldsymbol{\beta}_j+\boldsymbol{\epsilon}_j,\quad j=1,2,\cdots,L \tag{1.1}$$

$\mathbf{y}_j$ は方程式 j の従属変数の観測の $n\times1$ ベクトル，$\mathbf{X}_j$ はその方程式の K_j 個の説明変数の観測の $n\times K_j$ 行列，$\boldsymbol{\beta}_j$ はそれに対応する推定さるべきパラメーターの $K_j\times1$ ベクトル，$\boldsymbol{\epsilon}_j$ は攪乱項の $n\times1$ ベクトルである。(1)

このモデルにおける各方程式は標準の単1方程式モデルとして述べられているものである。しかしながら，それのすべての仮定をここで設けても，このモデルの確率的スペシフィケーションは完全ではない。単1方程式モデルにおいては，当然のことであるが，攪乱項の方程式間の相関について何も定めていない。$\boldsymbol{\epsilon}_j$ と $\boldsymbol{\epsilon}_l (j\neq l)$ とのあいだに相関があれば，方程式 j についての情報が方程式 l に含まれている。それで方程式 j をシステムの残余のものより切離して取扱うときには，残余のものに含まれている情報を利用しないことになる。

方程式の攪乱項は，1つ1つではとるに足らないために個別的に注目されなかった多くの影響を反映しているので，多くのモデルにおいて $\epsilon_{\alpha j}$ が $\epsilon_{\alpha l}$ とが（同期の）相関をもっと仮定することができる。

また，観測がランタムサンプルであるときには $\epsilon_{\alpha j}$ と $\epsilon_{\eta l} (\alpha\neq\eta)$ は異期無相関である。それで (1.1) についてつぎの仮定を設けておく。

$$\boldsymbol{\epsilon}_j\sim N(\mathbf{0},\sigma_{jj}\mathbf{I})$$

---

(1) Theil は joint generalized least-squares estimation technique としてこれを述べている。Kmenta, Gilbert (1968) はこれを最尤法として定式化している。

$$(1.2)\qquad E(\boldsymbol{\epsilon}_j\boldsymbol{\epsilon}_l')=\sigma_{jl}\mathbf{I}$$
$$E(\mathbf{X}_j'\boldsymbol{\epsilon}_l)=\mathbf{0}\quad j, l=1, 2, \cdots, L$$

このモデルを単1方程式モデルと同じように取扱うことを考慮してつぎの記号を準備しておく。

$$\mathbf{Y}=[\mathbf{y}_1\ \ \mathbf{y}_2\ \ \cdots\ \ \mathbf{y}_L]=\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1L} \\ y_{21} & y_{22} & \cdots & y_{2L} \\ \vdots & \vdots & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nL} \end{bmatrix}\qquad \mathbf{y}=\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_L \end{bmatrix}$$

$$(1.3)\qquad \mathbf{E}=[\boldsymbol{\epsilon}_1\ \ \boldsymbol{\epsilon}_2\ \ \cdots\ \ \boldsymbol{\epsilon}_L]=\begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{1L} \\ \epsilon_{21} & \epsilon_{22} & \cdots & \epsilon_{2L} \\ \vdots & \vdots & & \vdots \\ \epsilon_{n1} & \epsilon_{n2} & \cdots & \epsilon_{nL} \end{bmatrix}\qquad \boldsymbol{\epsilon}=\begin{bmatrix} \boldsymbol{\epsilon}_1 \\ \boldsymbol{\epsilon}_2 \\ \vdots \\ \boldsymbol{\epsilon}_L \end{bmatrix}$$

$$\overline{\mathbf{X}}=\begin{bmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_L \end{bmatrix}\qquad \boldsymbol{\beta}=\begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_L \end{bmatrix}$$

すなわち，$\mathbf{y}, \boldsymbol{\epsilon}$ はいずれも $nL\times 1$ ベクトル，$\overline{\mathbf{X}}$ は $(nL)\times(\Sigma K_j)$ 行列，$\boldsymbol{\beta}$ は $(\Sigma K_j)\times 1$ ベクトルである。この記号を用いるとモデル (1.1) をつぎのようにあらわすことができる。

$$(1.4)\qquad \mathbf{y}=\overline{\mathbf{X}}\boldsymbol{\beta}+\boldsymbol{\epsilon}$$

ここに $n\times 1$ の撹乱項ベクトルが L 個縦に配列されている。撹乱項の分散・共分散行列は，仮定 (1.2) を考慮するとつぎのようになる。

$$V(\boldsymbol{\epsilon})=E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}')$$
$$=E\begin{bmatrix} \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1' & \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2' & \cdots & \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_L' \\ \boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1' & \boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2' & \cdots & \boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_L' \\ \vdots & \vdots & & \vdots \\ \boldsymbol{\epsilon}_L\boldsymbol{\epsilon}_1' & \boldsymbol{\epsilon}_L\boldsymbol{\epsilon}_2' & \cdots & \boldsymbol{\epsilon}_L\boldsymbol{\epsilon}_L' \end{bmatrix}$$
$$=\begin{bmatrix} \sigma_{11}\mathbf{I} & \sigma_{12}\mathbf{I} & \cdots & \sigma_{1L}\mathbf{I} \\ \sigma_{21}\mathbf{I} & \sigma_{22}\mathbf{I} & \cdots & \sigma_{2L}\mathbf{I} \\ \vdots & \vdots & & \vdots \\ \sigma_{L1}\mathbf{I} & \sigma_{L2}\mathbf{I} & \cdots & \sigma_{LL}\mathbf{I} \end{bmatrix}$$
$$=\boldsymbol{\Sigma}\otimes\mathbf{I}_n$$

ここに⊗はクロネッカー積 Kronecker product である。[2] そして

$$(1.5)\qquad \mathbf{\Sigma}=\begin{bmatrix}\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1L}\\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2L}\\ \vdots & \vdots & & \vdots\\ \sigma_{L1} & \sigma_{L2} & \cdots & \sigma_{LL}\end{bmatrix}$$

である。このとき，$\boldsymbol{\epsilon}$ はつぎの結合密度関数をもつ。

$$(1.6)\qquad f(\boldsymbol{\epsilon})=(2\pi)^{-Ln/2}|(\mathbf{\Sigma}\otimes\mathbf{I})^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}\boldsymbol{\epsilon}'(\mathbf{\Sigma}\otimes)^{-1}\boldsymbol{\epsilon}$$

ここで変換 (1.4) を加えると，ヤコービアンが1であるので，$\mathbf{y}$ の結合密度関数，したがって尤度関数としてつぎのものをえる。

$$(1.7)\qquad l(\boldsymbol{\beta},\mathbf{\Sigma}\otimes\mathbf{I})=(2\pi)^{-Ln/2}|(\mathbf{\Sigma}\otimes\mathbf{I})^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}(\mathbf{y}-\overline{\mathbf{X}}\boldsymbol{\beta})'(\mathbf{\Sigma}\otimes\mathbf{I})^{-1}(\mathbf{y}-\overline{\mathbf{X}}\boldsymbol{\beta})$$

これの最大化によりつぎの推定量をえる。

$$(1.8)\qquad \tilde{\boldsymbol{\beta}}=[\overline{\mathbf{X}}'(\mathbf{\Sigma}\otimes\mathbf{I})^{-1}\overline{\mathbf{X}}]^{-1}\overline{\mathbf{X}}'(\mathbf{\Sigma}\otimes\mathbf{I})^{-1}\mathbf{y}$$
$$\tilde{\boldsymbol{\beta}}\sim N[\boldsymbol{\beta},\ \{\overline{\mathbf{X}}'(\mathbf{\Sigma}\otimes\mathbf{I})^{-1}\overline{\mathbf{X}}\}^{-1}]$$
$$\tilde{\mathbf{\Sigma}}=[\tilde{\sigma}_{jl}]$$
$$\tilde{\sigma}_{jl}=\frac{1}{n}(\mathbf{y}_j-\mathbf{X}_j\tilde{\boldsymbol{\beta}}_j)'(\mathbf{y}_l-\mathbf{X}_l\tilde{\boldsymbol{\beta}}_l)\qquad j,l=1,2,\cdots,L$$

観測行列の拡大

いま $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_L$ のそれぞれの代りにその和集合，$n\times K$ の行列 $\mathbf{X}$ を用い，(1.1) をつぎのように書きかえておく。

$$(1.9)\qquad \mathbf{y}_j=\mathbf{X}\boldsymbol{\beta}_j^*+\boldsymbol{\epsilon}_j\quad j=1,2,\cdots,L$$

(1.1) においては $\mathbf{X}_j$ について $K_j\times1$ の係数ベクトル $\boldsymbol{\beta}_j$ を用いたが，ここでは $\mathbf{X}$ についてそれに相応するように拡大された $K\times1$ の係数ベクトル $\boldsymbol{\beta}_j^*$ を用いる。[3]

(2)　クロネッカー積についてはたとえば Theil, p. 303 以下をみよ。

(3)　モデルの作成にあたってアプリオリ情報を考慮に入れない場合は星印つき記号を用い，それを考慮に入れるときには星印なしの記号を用いる。後者にはつぎの2つの場合がある。

横配列のとき，当該パラメーターを0とおく。例 (2.4)。

縦配列のとき，当該パラメーターをモデルよりはずす。例 (1.4)，(3.1)。

$$\mathbf{B}^* = [\boldsymbol{\beta}_1{}^* \quad \boldsymbol{\beta}_2{}^* \cdots \boldsymbol{\beta}_L{}^*]$$

$$= \begin{bmatrix} \beta_{11}{}^* & \beta_{12}{}^* \cdots & \beta_{1L}{}^* \\ \beta_{21}{}^* & \beta_{22}{}^* \cdots & \beta_{2L}{}^* \\ \vdots & \vdots & \vdots \\ \beta_{K1}{}^* & \beta_{K2}{}^* \cdots & \beta_{KL}{}^* \end{bmatrix} \qquad \boldsymbol{\beta}^* = \begin{bmatrix} \boldsymbol{\beta}_1{}^* \\ \boldsymbol{\beta}_2{}^* \\ \vdots \\ \boldsymbol{\beta}_L{}^* \end{bmatrix}$$

この $KL\times 1$ ベクトル $\boldsymbol{\beta}^*$ を用いると (1.9) はつぎのようにあらわされる。

(1.10) $$\mathbf{y} = (\mathbf{I}\otimes\mathbf{X})\boldsymbol{\beta}^* + \boldsymbol{\epsilon}$$

そして尤度関数はつぎのようになる。

(1.11) $$l(\boldsymbol{\beta}^*, \boldsymbol{\Sigma}\otimes\mathbf{I}) = (2\pi)^{-Ln/2}|(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}|^{1/2}$$
$$\exp -\frac{1}{2}[\mathbf{y}-(\mathbf{I}\otimes\mathbf{X})\boldsymbol{\beta}^*]'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}[\mathbf{y}-(\mathbf{I}\otimes\mathbf{X})\boldsymbol{\beta}^*]$$

これの最大化によってつぎの推定量をえる。

(1.12) $$\tilde{\boldsymbol{\beta}}^* = [\mathbf{I}\otimes(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y}$$
$$\tilde{\boldsymbol{\beta}}^* \sim N[\boldsymbol{\beta}^*, \boldsymbol{\Sigma}\otimes(\mathbf{X}'\mathbf{X})^{-1}]$$
$$\tilde{\boldsymbol{\Sigma}} = [\tilde{\sigma}_{jl}]$$
$$\tilde{\sigma}_{jl} = \frac{1}{n}(\mathbf{y}_j - \mathbf{X}\tilde{\boldsymbol{\beta}}_j{}^*)'(\mathbf{y}_l - \mathbf{X}\tilde{\boldsymbol{\beta}}_l{}^*)$$

ここに $\tilde{\boldsymbol{\beta}}^*$ は個々の方程式の係数の普通の最小2乗推定量を縦に配列したものであることがみられる。

ここでモデルの作成にあたって $\boldsymbol{\beta}^*$ に追加的に含まれた要素を0とおくとまえと同じ内容のモデルをえることができる。しかしそうしても一般に，その推定値として0をえることはできない。推定値として0値をえるには，尤度関数の最大化にあたってつぎの制約条件をつける。

$$0 = \mathbf{R}\boldsymbol{\beta}^*$$

ここに $\mathbf{R}$ は KL 次の単位行列において (i, i) 要素——i は $\boldsymbol{\beta}^*$ においてアプリオリに0である要素の番号をあらわす——を0としたものをあらわす。このときには複雑な形の推定量をえる。(4)

もとのシステム (1.4) へもどり，それに操作変数法——最大尤度法

(4) Theil, p. 285。

として定式化されているもの——を適用しよう。そのため (1.4) の左よりつぎの $Ln\times(\Sigma K_i)$ のブロック対角行列を掛けておく。

$$\mathbf{P}'=\mathrm{diag}(\mathbf{P}_1'\ \mathbf{P}_2'\cdots\mathbf{P}_L') \tag{1.13}$$

ここに $\mathbf{P}_j$ は操作変数の $n\times K_j$ の観測行列をあらわす。$(j=1,2,\cdots,L)$ このとき変換後の攪乱項はつぎの結合密度関数をもつ。

$$f(\mathbf{P}'\boldsymbol{\epsilon})=(2\pi)^{-Ln/2}|[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{P}]^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}(\mathbf{P}'\boldsymbol{\epsilon})'[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{P}]^{-1}\mathbf{P}'\boldsymbol{\epsilon} \tag{1.14}$$

ここで変換

$$\mathbf{P}'\mathbf{y}=\mathbf{P}'\mathbf{X}\boldsymbol{\beta}+\mathbf{P}'\boldsymbol{\epsilon}$$

を加えると，ヤコービアンが1であるので，つぎの尤度関数をえる。

$$l[\boldsymbol{\beta},\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{P}]=(2\pi)^{-Ln/2}|[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{P}]^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}[\mathbf{P}'\mathbf{y}-\mathbf{P}'\mathbf{X}\boldsymbol{\beta}]'[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{P}]^{-1}[\mathbf{P}'\mathbf{y}-\mathbf{P}'\mathbf{X}\boldsymbol{\beta}] \tag{1.15}$$

これの最大化によってつぎの推定量をえる。

$$\begin{aligned}\tilde{\boldsymbol{\beta}}&=\{\mathbf{X}'\mathbf{P}[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}\mathbf{P}]^{-1}\mathbf{P}'\mathbf{X}\}^{-1}\mathbf{X}'\mathbf{P}[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}\mathbf{P}]^{-1}\mathbf{P}'\mathbf{y}\\&=(\mathbf{P}'\mathbf{X})^{-1}\{\mathbf{X}'\mathbf{P}[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}\mathbf{P}]^{-1}\}^{-1}\mathbf{X}'\mathbf{P}[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}\mathbf{P}]^{-1}\mathbf{P}'\mathbf{y}\\&=(\mathbf{P}'\mathbf{X})^{-1}\mathbf{P}'\mathbf{y}\end{aligned} \tag{1.16}$$

$$\tilde{\boldsymbol{\beta}}\sim N[\boldsymbol{\beta},\ (\mathbf{P}'\mathbf{X})^{-1}\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{P}[(\mathbf{P}'\mathbf{X})^{-1}]']$$

これが操作変数法をシステム (1.4) へ適用してえる推定量である。

攪乱項ベクトルの横配列

攪乱項ベクトルを横に配列して (1.9) をつぎのようにあらわす。

$$\begin{aligned}\mathbf{E}&=[\boldsymbol{\epsilon}_1\ \boldsymbol{\epsilon}_2\cdots\boldsymbol{\epsilon}_L]\\&=[\mathbf{y}_1-\mathbf{X}\boldsymbol{\beta}_1^*\quad \mathbf{y}_2-\mathbf{X}\boldsymbol{\beta}_2^*\cdots\mathbf{y}_L-\mathbf{X}\boldsymbol{\beta}_L^*]\\&=\mathbf{Y}-\mathbf{X}\mathbf{B}^*\end{aligned} \tag{1.17}$$

この攪乱項の結合密度はつぎのようにあらわされる。

$$f(\mathbf{E})=(2\pi)^{-Ln/2}|\boldsymbol{\Sigma}^{-1}|^{n/2}\exp-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}^{-1}\mathbf{E}'\mathbf{E} \tag{1.18}$$

これに変換 (1.17) を加えるとつぎの尤度関数をえる。

$$l(\mathbf{B}^*,\boldsymbol{\Sigma})=(2\pi)^{-Ln/2}|\boldsymbol{\Sigma}^{-1}|^{n/2}$$
$$\exp-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}^{-1}[\mathbf{Y}-\mathbf{X}\mathbf{B}^*]'[\mathbf{Y}-\mathbf{X}\mathbf{B}^*] \tag{1.19}$$

これと (1.11) との同等性を示すことができる。これの最大化によってつぎの最尤推定量をえる。[5]

$$\tilde{\mathbf{B}}^* = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$
$$\tilde{\mathbf{\Sigma}} = \frac{1}{n}[\mathbf{Y} - \mathbf{X}\tilde{\mathbf{B}}^*]'[\mathbf{Y} - \mathbf{X}\tilde{\mathbf{B}}^*]$$

縦配列と横配列との関係

$L \times n$ の行列を $\mathbf{S}$, $n \times L$ の行列を $\mathbf{E}$ とするとき

$$\mathrm{tr}(\mathbf{SE}) = \sum_j \sum_j s_{ij}\epsilon_{ji}$$
$$\mathrm{tr}(\mathbf{ES}) = \sum_j \sum_i \epsilon_{ji} s_{ij}$$

であり，つぎの関係が成立する。

$$\mathrm{tr}(\mathbf{SE}) = \mathrm{tr}(\mathbf{ES})$$

いま $n \times L$ の行列を $\mathbf{E}$, $L \times L$ の行列を $\mathbf{\Sigma}^{-1} = [\sigma^{ij}]$ とすると $\mathbf{\Sigma}^{-1}\mathbf{E}'$ は $L \times n$ の行列, $\mathbf{E}$ は $n \times L$ の行列であるから,

$$\mathrm{tr}\mathbf{E}\mathbf{\Sigma}^{-1}\mathbf{E}' = \mathrm{tr}\mathbf{\Sigma}^{-1}\mathbf{E}'\mathbf{E}$$

をえる。

つぎに (1.11) と (1.18) が同じであること，とくにその指数部分が同じであることを示しておこう。ここでつぎの関係よりはじめる。

$$\mathbf{\Sigma}^{-1}\mathbf{E}' = \begin{bmatrix} \sum \sigma^{1i}\epsilon_{1i} & \cdots & \sum \sigma^{1i}\epsilon_{ni} \\ \vdots & & \vdots \\ \sum \sigma^{Li}\epsilon_{1i} & \cdots & \sum \sigma^{Li}\epsilon_{ni} \end{bmatrix}$$

$$\mathbf{\Sigma}^{-1}\mathbf{E}'\mathbf{E} = \begin{bmatrix} \sum_\alpha \epsilon_{\alpha 1} \sum_i \sigma^{1i}\epsilon_{\alpha i} & \cdots & \sum_\alpha \varepsilon_{\alpha L} \sum_i \sigma^{1i}\epsilon_{\alpha i} \\ \vdots & & \vdots \\ \sum_\alpha \epsilon_{\alpha L} \sum \sigma^{Li}\varepsilon_{\alpha i} & \cdots & \sum \epsilon_{\alpha L} \sum_\alpha \sigma^{Li}\epsilon_{\alpha i} \end{bmatrix}$$

$$\begin{aligned}
\mathrm{tr}\mathbf{\Sigma}^{-1}\mathbf{E}'\mathbf{E} &= \sum_\alpha \epsilon_{\alpha 1} \sum_i \sigma^{1i}\epsilon_{\alpha i} + \cdots + \sum_\alpha \epsilon_{\alpha L} \sum \sigma^{Li}\epsilon_{\alpha i} \\
&= \begin{Bmatrix} \epsilon_{11} \sum \sigma^{1i}\epsilon_{1i} \\ + \cdots \\ + \epsilon_{ni} \sum \sigma^{1i}\epsilon_{ni} \end{Bmatrix} + \cdots + \begin{Bmatrix} \epsilon_{1L} \sum \sigma^{L1}\epsilon_{1i} \\ + \cdots \\ + \epsilon_{nL} \sum \sigma^{Li}\epsilon_{ni} \end{Bmatrix} \\
&= \begin{bmatrix} \epsilon_{11} \\ \vdots \\ \epsilon_{n1} \end{bmatrix}' \begin{bmatrix} \sum \sigma^{1i}\epsilon_{1i} \\ \vdots \\ \sum \sigma^{1i}\epsilon_{ni} \end{bmatrix} + \cdots + \begin{bmatrix} \epsilon_{1L} \\ \vdots \\ \epsilon_{nL} \end{bmatrix}' \begin{bmatrix} \sum \sigma^{Li}\epsilon_{1i} \\ \vdots \\ \sum \sigma^{Li}\epsilon_{ni} \end{bmatrix} \\
&= \boldsymbol{\epsilon}_1'(\sigma^{11}\boldsymbol{\epsilon}_1 + \cdots + \sigma^{1L}\boldsymbol{\epsilon}_L) + \cdots + \boldsymbol{\epsilon}_L'(\sigma^{L1}\boldsymbol{\epsilon}_1 + \cdots + \sigma^{LL}\boldsymbol{\epsilon}_L)
\end{aligned}$$

---

(5) Theil, p. 501 を参照のこと。

$$=\begin{bmatrix}\boldsymbol{\epsilon}_1\\ \vdots\\ \boldsymbol{\epsilon}_L\end{bmatrix}'\begin{bmatrix}\sigma^{11}\mathbf{I} & \cdots & \sigma^{1L}\mathbf{I}\\ & \cdots & \\ \sigma^{L1}\mathbf{I} & \cdots & \sigma^{LL}\mathbf{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{\epsilon}_1\\ \vdots\\ \boldsymbol{\epsilon}_L\end{bmatrix}$$

$$=\boldsymbol{\epsilon}'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I})\boldsymbol{\epsilon}$$

したがってつぎの関係をえる。

(1.20)　$$\mathrm{tr}\boldsymbol{\Sigma}^{-1}\mathbf{E}'\mathbf{E}=\mathrm{tr}\mathbf{E}\boldsymbol{\Sigma}^{-1}\mathbf{E}'=\boldsymbol{\epsilon}'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I})\boldsymbol{\epsilon}$$

これに対応して，尤度関数の指数部分についてつぎの関係をえる。

(1.21)　$$[\mathbf{y}-(\mathbf{I}\otimes\mathbf{X})\boldsymbol{\beta}^*]'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}[\mathbf{y}-(\mathbf{I}\otimes\mathbf{X})\boldsymbol{\beta}^*]$$
$$=\mathrm{tr}\boldsymbol{\Sigma}^{-1}[\mathbf{Y}-\mathbf{X}\mathbf{B}^*]'[\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}^*]$$

左辺は（1.11）のもの，右辺は（1.18）のものである。これよりつぎのものをえる。

(1.22)　$$\min[\mathbf{y}-(\mathbf{I}\otimes\mathbf{X})\boldsymbol{\beta}^*]'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}[\mathbf{y}-(\mathbf{I}\otimes\mathbf{X})\boldsymbol{\beta}^*]$$
$$=\min\ \mathrm{tr}\,\boldsymbol{\Sigma}^{-1}[\mathbf{Y}-\mathbf{X}\mathbf{B}^*]'[\mathbf{Y}-\mathbf{X}\mathbf{B}^*]$$

第2節　連立方程式モデル　攪乱項ベクトルの横配列

完全情報量尤法

結合縦属変数の変動を説明するために，つぎの連立方程式モデルを用いる。そしてこれが経済の構造方程式システムとよばれている。

(2.1)　$$\mathbf{y}_j=\mathbf{Y}\mathbf{c}_j^*+\mathbf{X}\boldsymbol{\beta}_j^*+\boldsymbol{\epsilon}_j\quad j=1,2,\cdots,L$$

$\mathbf{Y}$ はすべての結合縦属変数の $n\times L$ 観測行列，$\mathbf{y}_j$ は j 番目の結合縦属変数で，方程式 j において（規準化より）係数1をもつものの $n\times 1$ 観測ベクトルである。$\mathbf{X}$ はすべての先決変数の $n\times K$ 観測行列である。そして

$$\mathbf{c}_j^*=[c_{1j}^*\ c_{2j}^*\cdots c_{Lj}^*]',\quad c_{jj}^*=0$$

は $L\times1$ の係数ベクトル，$\boldsymbol{\beta}_j^*$ は $K\times1$ の係数ベクトルである。ここではアプリオリ情報を考慮に入れていないことを星印をつけて示している。

攪乱項ベクトルを横に配列してモデルをつぎのようにあらわしておく。

(2.2)
$$\mathbf{Y}=\mathbf{Y}\mathbf{C}^*+\mathbf{X}\mathbf{B}^*+\mathbf{E}$$

ここに $\mathbf{E}$ は $n\times L$ の攪乱項の行列である。(1.3) をみよ。そして

$$\mathbf{C}^*=[\mathbf{c}_1{}^*\ \mathbf{c}_2{}^*\ \cdots\ \mathbf{c}_L{}^*]$$
$$\mathbf{B}^*=[\boldsymbol{\beta}_1{}^*\ \boldsymbol{\beta}_2{}^*\ \cdots\ \boldsymbol{\beta}_L{}^*]$$

はそれぞれの $L\times L$, $K\times L$ 係数行列をあらわす。連立方程式モデル (2.2) を前節の多数方程式モデル (1.17) と比較すると，$\mathbf{YC}^*$ の項が追加されていることがみられる。(6) そしてこれが被説明変数と同じ結合従属変数のものであることがこのモデルの分析を複雑にしている最大の要因である。

この攪乱項が (1.2) と同じ仮定を満足しているものとする。攪乱項 $\mathbf{E}$ の結合密度関数は (1.18) であらわされている。ここで尤度関数をえるための変換式は (2.2) である。これを用いるとつぎのものをえる。

(2.3)
$$l(\mathbf{C}^*,\mathbf{B}^*,\boldsymbol{\Sigma})=(|\mathbf{I}-\mathbf{C}^*|^2)^{n/2}(2\pi)^{-Ln/2}|\boldsymbol{\Sigma}^{-1}|^{n/2}$$
$$\exp-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}^{-1}[\mathbf{Y}(\mathbf{I}-\mathbf{C}^*)-\mathbf{X}\mathbf{B}^*]'[\mathbf{Y}(\mathbf{I}-\mathbf{C}^*)-\mathbf{X}\mathbf{B}^*]$$

これを最大化するものが完全情報最尤法である。(7) ここには構造パラメーターがヤコービアンを経由して非線型の形で入っているために，それの完全情報最尤推定量を明示的な式であらわすことはできない。

間接最小2乗法

連立方程式モデルの推定上の複雑性を迂回するのに1つの巧妙な方法がある。それを説明するためにいまアプリオリにないことのわかっている構造パラメーターを0とおき，構造方程式システム (2.2) をつぎのように星印を除いた記号 $\mathbf{C}, \mathbf{B}$ を用いてあらわしておく。

(2.4)
$$\mathbf{Y}=\mathbf{Y}\mathbf{C}+\mathbf{X}\mathbf{B}+\mathbf{E}$$

$\mathbf{I}-\mathbf{C}$ の非特異性を仮定して，この誘導型システムをつぎのように

---

(6) あるいは(2.2)を $\mathbf{Y}(\mathbf{I}-\mathbf{C}^*)-\mathbf{X}\mathbf{B}^*=\mathbf{E}$ と書き改めて，(1.17) $\mathbf{Y}-\mathbf{X}\mathbf{B}^*=\mathbf{E}$ と比較すると $\mathbf{Y}$ の係数行列が $\mathbf{I}$ より $\mathbf{I}-\mathbf{C}^*$ に改められていることがみられる。

(7) Koopmans, Rubin, Leipnik, (1950)。

あらわしておく。[8]

$$
\begin{aligned}
&\mathbf{Y}=\mathbf{X\Pi}+\mathbf{V}\\
(2.5)\qquad &\mathbf{\Pi}=\mathbf{B}(\mathbf{I}-\mathbf{C})^{-1}\\
&\mathbf{V}=\mathbf{E}(\mathbf{I}-\mathbf{C})^{-1}
\end{aligned}
$$

$\mathbf{V}, \mathbf{E}$ の行 t を $\mathbf{V}_t, \mathbf{E}_t$ とあらわすと

$$\mathbf{V}_t=\mathbf{E}_t(\mathbf{I}-\mathbf{C})^{-1}$$

であることはいうまでもない。$\mathbf{E}$ について設けた仮定より $\mathbf{V}$ はつぎの性質をもつ,

$$
\begin{aligned}
&\mathbf{V}_t'\sim N(\mathbf{0}, \mathbf{\Omega})\\
&\mathbf{\Omega}=[(\mathbf{I}-\mathbf{C})^{-1}]'\mathbf{\Sigma}(\mathbf{I}-\mathbf{C})^{-1}\\
&E[\mathbf{V}_t'\mathbf{V}_s]=\delta_{ts}\mathbf{\Omega},\ \ t, s=1, 2, \cdots, n\\
&E[\mathbf{X}'\mathbf{V}]=\mathbf{0}
\end{aligned}
$$

この $\mathbf{V}$ の結合密度関数をつぎのようにあらわしておく。

$$f(\mathbf{V})=(2\pi)^{-Ln/2}|\mathbf{\Omega}^{-1}|^{n/2}\exp-\frac{1}{2}\mathrm{tr}\mathbf{\Omega}^{-1}\mathbf{V}'\mathbf{V}$$

これに変換 (2.4) を加える。このヤコービアンが 1 であるので, $\mathbf{Y}$ の結合密度関数，したがって尤度関数としてつぎのものをえる。

$$
\begin{aligned}
(2.6)\qquad l(\mathbf{\Pi}, \mathbf{\Omega})=&(2\pi)^{-Ln/2}|\mathbf{\Omega}^{-1}|^{n/2}\\
&\exp-\frac{1}{2}\mathrm{tr}\mathbf{\Omega}^{-1}[\mathbf{Y}-\mathbf{X\Pi}]'[\mathbf{Y}-\mathbf{X\Pi}]
\end{aligned}
$$

これの最大化によりつぎの推定量をえる。

$$
\begin{aligned}
(2.7)\qquad &\tilde{\mathbf{\Pi}}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\\
&\tilde{\mathbf{\Omega}}=\frac{1}{n}[\mathbf{Y}-\mathbf{X}\tilde{\mathbf{\Pi}}]'[\mathbf{Y}-\mathbf{X}\tilde{\mathbf{\Pi}}]
\end{aligned}
$$

(2.5) において (2.7) を用いてえる連立方程式

$$(2.8)\qquad \tilde{\mathbf{\Pi}}(\mathbf{I}-\mathbf{C})=\mathbf{B}$$

を $\mathbf{C}, \mathbf{B}$ について解いてその推定量をえる。これが間接最小 2 乗法 Indirect least squares とよばれているものに等しい。

識別の条件

---

(8)　システム (2.2) についても同じように誘導型をもとめることができる。Theil, p. 441。

(2.8) を $\mathbf{C}, \mathbf{B}$ について解くことができるためにはある条件がみたされていなければならない。この条件を (2.4) の第1の構造方程式((2.1) について述べると，$j=1$ とおき星印を除いた場合)

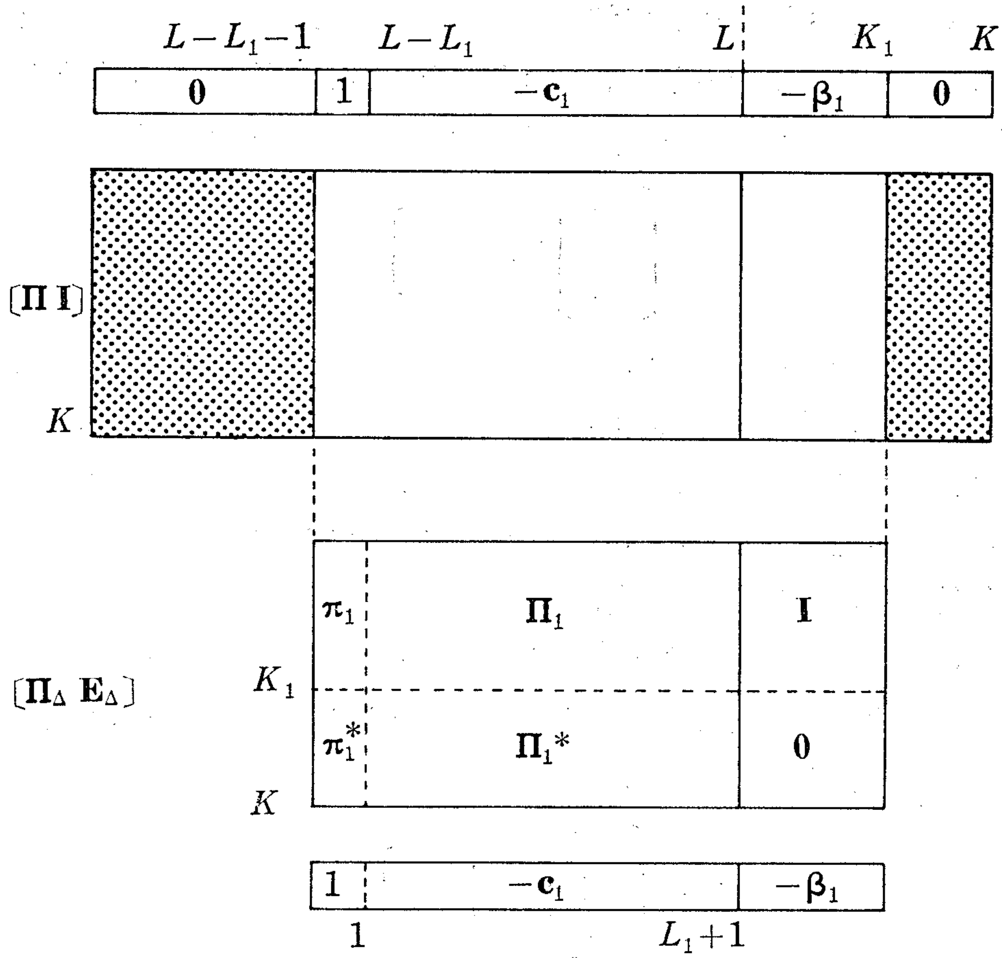
$$\epsilon_1=\mathbf{Y}\begin{bmatrix}1\\-\mathbf{c}_1\\\mathbf{0}\end{bmatrix}+\mathbf{X}\begin{bmatrix}-\boldsymbol{\beta}_1\\\mathbf{0}\end{bmatrix} \tag{2.9}$$

の係数についてもとめておこう。ここに係数ベクトルは (2.4) の係数行列 $\mathbf{I}-\mathbf{C}, \mathbf{B}$ それぞれの列1であり，$L\times1$, $K\times1$ のものである。ここにアプリオリに0とわかっている要素は (必要であれば適当に順序を入れかえて) うしろへまとめておく。第1項の $\mathbf{0}$ は $(L-L_1-1)\times1$, 第2項の $\mathbf{0}$ は $(K-K_1)\times1$ のものである。

この係数ベクトルの推定のためにわれわれは式 (2.8) をつかっている。この式のうちわれわれの注目している係数ベクトルを含むものをつぎのようにあらわしておく。

$$\begin{aligned}0&=\boldsymbol{\Pi}\begin{bmatrix}\mathbf{0}\\1\\-\mathbf{c}_1\end{bmatrix}+\mathbf{I}_K\begin{bmatrix}-\boldsymbol{\beta}_1\\\mathbf{0}\end{bmatrix}\\&=\boldsymbol{\Pi}_\Delta\begin{bmatrix}1\\-\mathbf{c}_1\end{bmatrix}+\mathbf{E}_\Delta(-\boldsymbol{\beta}_1)\\&=[\boldsymbol{\Pi}_\Delta\ \ \mathbf{E}_\Delta]\begin{bmatrix}1\\-\mathbf{c}_1\\-\boldsymbol{\beta}_1\end{bmatrix}\\&=\begin{bmatrix}\boldsymbol{\pi}_1 & \boldsymbol{\Pi}_1 & \mathbf{I}\\\boldsymbol{\pi}_1^* & \boldsymbol{\Pi}_1^* & \mathbf{0}\end{bmatrix}\begin{bmatrix}1\\-\mathbf{c}_1\\-\boldsymbol{\beta}_1\end{bmatrix}\end{aligned} \tag{2.10}$$

ここに $\mathbf{E}'_\Delta=[\mathbf{I}\ \mathbf{0}]$ ($\mathbf{I}$ は K_1 次の単位行列) であり，$\boldsymbol{\Pi}_\Delta$ は $\boldsymbol{\Pi}$ のうしろの L_1+1 個の列よりなる部分行列である。なお，最後の式においては $\boldsymbol{\Pi}_\Delta$ を1個の列，L_1 個の列および K_1 個の行，$K^*=K-K_1$ 個の行に分割して示しておいた。この最初の式と最後の式を図に示しておいた。



(2.7) でえた推定量 $\tilde{\mathbf{\Pi}}$ を代入して (2.10) を $\mathbf{c}_1, \boldsymbol{\beta}_1$ の推定量について解く。(2.10) は継起システムであり，後半のものより $\mathbf{c}_1$ をもとめ，前半のものより $\mathbf{c}_1$ を所与として $\boldsymbol{\beta}_1$ をもとめる。後半のシステム

$$\mathbf{\Pi}_1^*\mathbf{c}_1=\boldsymbol{\pi}_1^*$$

が $\mathbf{c}_1$ について一意の解をもつための必要かつ十分な条件が

$$\mathrm{rank}\,(\mathbf{\Pi}_1^*)=L_1$$

である。これが第1の構造方程式の識別のランク条件 rank condition for identification, 適度識別の条件 condition of just identification である。

これが成立するための1つの必要条件が

$$K \geq K_1+L_1$$

であることは明らかである。これが識別のオーダー条件 order condition for identification である。

なお，タイルは識別の条件を述べているが間接最小2乗法をとりあげることはしていない。識別のランク条件（オーダー条件さえ）を満足するモデルが稀有と考えるからであろう。

制限情報最尤法

$\mathbf{\Pi}_1{}^*$ のランクが L_1 に足らないときには構造パラメーターを推定することはできない。ではそのランクが L_1 を越え過剰識別のときにはどうするか。多くのモデルにおいてこうなることがみられるのでこの場合が処理できなければならない。

注目している第1の構造方程式にあらわれている（0係数をもたないとアプリオリに考えられている）L_1+1 個の結合従属変数の観測行列 Y_Δ の尤度関数を用いる。これはつぎの点を修正すると (2.6) と同じである。$\mathbf{Y}$ を $\mathbf{Y}_\Delta$ で，$\mathbf{\Pi}$ を $\mathbf{\Pi}_\Delta$ でおきかえる。Lを L_1+1 で，$\mathbf{\Omega}$ を $\mathbf{Y}_\Delta$ に対応する $(L_1+1)\times(L_1+1)$ の首座部分行列 $\mathbf{\Omega}_\Delta$ でおきかえる。すなわちつぎのようにあらわされる尤度関数を用いる。

$$l(\mathbf{\Pi}_\Delta, \mathbf{\Omega}_\Delta) = (2\pi)^{-(L+1)n/2}|\mathbf{\Omega}_\Delta{}^{-1}|^{n/2}$$
$$\exp -\frac{1}{2}\mathrm{tr}\mathbf{\Omega}_\Delta{}^{-1}[\mathbf{Y}_\Delta-\mathbf{X\Pi}_\Delta]'[\mathbf{Y}_\Delta-\mathbf{X\Pi}_\Delta]$$

そしてこれの最大化を制約条件 (2.13) のもとでする。これが制限情報最尤法に等しい。(9)

ここで用いた横配列の型についてタイルはつぎのように批判的考えを述べている。(10)

「記号 (3.6)〔われわれの (2.2) に相当する〕はまとまりがありエレガントである。そして誘導型 (3.8) にとって非常に有益である。しかし1つの構造方程式のパラメーターを推定しようとするときには本当に便利であるものではない。その理由は，パラメーター行列 $\mathbf{B}$, $\mathbf{\Gamma}$〔われわれの $\mathbf{I}-\mathbf{C}^*, \mathbf{B}^*$〕が多くの無駄を含んでいるからである。普

(9) Theil, pp. 678–686。
(10) Theil, pp. 443–444。

通，その多くの要素が 1, 0，あるいは－1であることがわかっているからである。これは (3.7) よりただちに明らかである。そして事実，問題 3.2 において $\mathbf{B}, \mathbf{\Gamma}$ の未知の要素の割合がクインのモデル1の場合には7分の1，ケインズ・モデル (1.1)，(1.2) の場合には6分の1であることを示す。したがって推定に先立って記号の問題について考える。[11]」

ここで横配列においてアプリオリにないことのわかっている係数は0とおいて処理することができるけれども，係数の推定へと進むとき，0推定値をえるためには制約条件つき最大尤度法を用いなければならない。これはつぎの縦配列によるほうが簡単に処理できる。

さらに縦配列の型を用いるときには過剰識別の場合を処理するのに特別の配慮は必要でなくなるという便もある。

第3節　連立方程式モデル　攪乱項ベクトルの縦配列

完全情報最尤法

ここでは攪乱項ベクトルを縦に配列したものに注目する。そしてアプリオリ情報を考慮後のものについて述べる。連立方程式モデル (2.1) をつぎのようにあらわしておく。

$$\mathbf{y}_j=\mathbf{Y}_j\mathbf{c}_j+\mathbf{X}_j\boldsymbol{\beta}_j+\boldsymbol{\epsilon}_j$$

$$(3.1)\qquad =\mathbf{Z}_j\boldsymbol{\delta}_j+\boldsymbol{\epsilon}_j$$

$$(3.2)\qquad \mathbf{Z}_j=[\mathbf{Y}_j\ \mathbf{X}_j]$$

$$\boldsymbol{\delta}_j=\begin{bmatrix}\mathbf{c}_j\\ \boldsymbol{\beta}_j\end{bmatrix},\quad j=1,2,\cdots,L$$

$\mathbf{Y}_j$ は結合従属変数の $n\times L_j$ 観測行列，$\mathbf{X}_j$ は先決変数の $n\times K_j$ の観測行列，$\mathbf{Z}_j$ は説明変数の $n\times(L_j+K_j)$ 観測行列をあらわす。$\mathbf{c}_j$, $\boldsymbol{\beta}_j$, $\boldsymbol{\delta}_j$ はそれぞれ $L_j\times 1$, $K_j\times 1$ および $(L_j+K_j)\times 1$ の係数ベクトルである。ここではアプリオリにないことのわかっているものを除いて次

(11)　Theil, pp. 443–444。

数をさけている点が前節のあらわし方と異なっている。$\mathbf{y}_j, \boldsymbol{\epsilon}_j$ は前節と同じものをあらわす。

攪乱項ベクトルを縦配列してこのモデルをつぎのようにあらわしておく。

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_L \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} \cdots \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 \cdots \mathbf{0} \\ \vdots & \vdots \quad\ \vdots \\ \mathbf{0} & \mathbf{0} \cdots \mathbf{Z}_L \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \\ \vdots \\ \boldsymbol{\delta}_L \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_1 \\ \boldsymbol{\epsilon}_2 \\ \vdots \\ \boldsymbol{\epsilon}_L \end{bmatrix}$$

$$\text{(3. 3)} \qquad \mathbf{y} = \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}$$

ここに $\mathbf{y}, \mathbf{Z}$ はそれぞれ $nL \times 1, nL \times \sum(L_j + K_j)$ の被説明変数，説明変数の観測行列，$\boldsymbol{\epsilon}$ は $nL \times 1$ の攪乱項のベクトル，$\boldsymbol{\delta}$ は $[\sum(L_j + K_j)] \times 1$ の係数ベクトルをあらわす。（横配列のときの $(K+L) \times L$ 個より $\sum(L_j + K_j)$ へ減っている。）この攪乱項は (1. 2) の性質をもつものとする。このときこの結合密度関数は (1. 6) と同じ形であらわされる。

ここでそれに変換 (3. 3) を加える。この変換が $\boldsymbol{\epsilon}$ より（$\mathbf{Z}$ を所与としての）$\mathbf{y}$ へのものであるとみてはならない。(3. 1) の第1式において明らかなように，$\boldsymbol{\epsilon}$ を（$\mathbf{X}$ を所与として）$\mathbf{Y}$ へ変換するものである。したがって，この変換のヤコービアンは1でなく $|(\mathbf{I}-\mathbf{C})^2|^{1/2}$ である。それでモデル (3. 3) のもとでの尤度関数としてつぎのものをえる。

$$l(\boldsymbol{\delta}, \boldsymbol{\Sigma}\otimes\mathbf{I}) = |(\mathbf{I}-\mathbf{C})^2|^{n/2}(2\pi)^{-Ln/2}|(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}|^{1/2}$$
$$\exp -\frac{1}{2}(\mathbf{y}-\mathbf{Z}\boldsymbol{\delta})'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}(\mathbf{y}-\mathbf{Z}\boldsymbol{\delta})$$

これをパラメーターに関して最大化するものが完全情報最尤法である。

前節で述べたように連立方程式モデルの最尤推定法を複雑にしている根源は非1のヤコービアンにある。以下においてはこれを無視することによって単純化する。縦配列のモデルについて述べられている推定方法はすべて（暗黙裡に）この仮定を用いているものと解釈できる。

参考

アプリオリ情報を考慮しないときには連立方程式モデル (2. 1) をつ

ぎのようにあらわすことができる。

$$\mathbf{y}_j=\mathbf{Y}\mathbf{c}_j{}^*+\mathbf{X}\boldsymbol{\beta}_j{}^*+\boldsymbol{\epsilon}_j$$
$$=\mathbf{W}\boldsymbol{\delta}_j{}^*+\boldsymbol{\epsilon}_j$$
$$\mathbf{W}=[\mathbf{Y}\ \mathbf{X}]$$
$$\boldsymbol{\delta}_j{}^*=\begin{bmatrix}\mathbf{c}_j{}^*\\ \boldsymbol{\beta}_j{}^*\end{bmatrix},\quad j=1,2,\cdots,L$$

ここではこれをつぎのように縦に配列しておく。

$$\mathbf{y}=(\mathbf{I}\otimes\mathbf{W})\boldsymbol{\delta}^*+\boldsymbol{\epsilon}$$

これよりつぎの尤度関数をえることができる。

$$(3.4)\qquad l(\boldsymbol{\delta}^*,\boldsymbol{\Sigma}\otimes\mathbf{I})=|(\mathbf{I}-\mathbf{C}^*)^2|^{n/2}(2\pi)^{-Ln/2}|(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}[\mathbf{y}-(\mathbf{I}\otimes\mathbf{W})\boldsymbol{\delta}^*](\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}[\mathbf{y}-(\mathbf{I}\otimes\mathbf{W})\boldsymbol{\delta}^*]$$

これがクープマンス等の用いた尤度関数 (2.3) と同じものである。これは多くの無默な係数を含んでいるだけでなく，アプリオリ情報を考慮に入れるために（ゼロ）制約条件つき最大尤度法を用いるときには推定方法は複雑なものになる。

操作変数法

ここで連立方程式モデル (3.3) へ操作変数法を適用しよう。そのため (3.3) の左より (1.22) の $\mathbf{P}'$ を掛ける。ただし，ここにその部分行列 $\mathbf{P}_j$ は $n\times(L_j+K_j)$ の操作変数の観測行列をあらわす。$j=1,2,\cdots,L$。

$$(3.5)\qquad \mathbf{P}'\mathbf{y}=\mathbf{P}'\mathbf{Z}\boldsymbol{\delta}+\mathbf{P}'\boldsymbol{\epsilon}$$

新しい攪乱項の結合密度関数は (1.14) と同じ形である。

これに変換 (3.8) を加えてつぎの尤度関数をえる。

$$(3.6)\qquad l[\boldsymbol{\delta},\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{P}]=|(\mathbf{I}-\mathbf{C})^2|^{n/2}(2\pi)^{-Ln/2}|[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{P}]^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}[\mathbf{P}'\mathbf{y}-\mathbf{P}'\mathbf{Z}\boldsymbol{\delta}]'[\mathbf{P}'(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{P}]^{-1}[\mathbf{P}'\mathbf{y}-\mathbf{P}'\mathbf{Z}\boldsymbol{\delta}]$$

ここでヤコービアンを無視したうえでこれを最大化してつぎの推定量をえる。

$$(3.7)\qquad \tilde{\boldsymbol{\delta}}=(\mathbf{P}'\mathbf{Z})^{-1}\mathbf{P}'\mathbf{y}$$

これが操作変数推定量である。(1.25) と比較してみよ。$\mathbf{P}, \mathbf{Z}$ がどちらもブロック対角行列であるから，これは各方程式に個別に操作変数法を適用してえるものと同じである。

操作変数の候補物はつぎの性質をもっている。操作変数はそれによって代られる変数と相関をもつべきであるという条件を考慮するときには，操作変数はモデルに現実にあらわれている変数（あるいはそれの関数）でなければならない。ある変数が候補物に値する相関をもつときには，このような変数はよいモデルにおいては用いられているはずであるからである。

また，操作変数は攪乱項と無相関であるという条件を考慮すると，候補物は結合従属変数を含むことはできない。それで候補物としては先決変数（あるいはその関数）のみが残る。それで操作変数の候補物はつぎのようにあらわされる。

$$\mathbf{P}=\mathbf{X}\mathbf{A} \tag{3.8}$$

3 段階最小 2 乗法

ここで $\mathbf{P}'$ として $\mathbf{I}\otimes\mathbf{X}'$ を用いる。ここに $\mathbf{X}$ はシステムのすべての先決変数の $n\times K$ 行列でてる。[(12)]

$$(\mathbf{I}\otimes\mathbf{X}')\mathbf{y}=(\mathbf{I}\otimes\mathbf{X}')\mathbf{Z}\boldsymbol{\delta}+(\mathbf{I}\otimes\mathbf{X}')\boldsymbol{\epsilon} \tag{3.9}$$

変換後の攪乱項の結合密度関数はつぎのようにあらわされる。

$$f[(\mathbf{I}\otimes\mathbf{X}')\boldsymbol{\epsilon}]=(2\pi)^{-Ln/2}|(\boldsymbol{\Sigma}\otimes\mathbf{X}'\mathbf{X})^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}[(\mathbf{I}\otimes\mathbf{X}')\boldsymbol{\epsilon}]'(\boldsymbol{\Sigma}\otimes\mathbf{X}'\mathbf{X})^{-1}[(\mathbf{I}\otimes\mathbf{X}')\boldsymbol{\epsilon}]$$

これに変換 (3.9) を加えるとつぎの尤度関数をえる。

$$l(\boldsymbol{\delta},\boldsymbol{\Sigma}\otimes\mathbf{X}'\mathbf{X})=|(\mathbf{I}-\mathbf{C})^2|^{n/2}(2\pi)^{-Ln/2}|(\boldsymbol{\Sigma}\otimes\mathbf{X}'\mathbf{X})^{-1}|^{1/2} \tag{3.10}$$
$$\exp-\frac{1}{2}[(\mathbf{I}\otimes\mathbf{X}')\mathbf{y}-(\mathbf{I}\otimes\mathbf{X}')\mathbf{Z}\boldsymbol{\delta}]'(\boldsymbol{\Sigma}\otimes\mathbf{X}'\mathbf{X})^{-1}$$
$$[(\mathbf{I}\otimes\mathbf{X}')\mathbf{y}-(\mathbf{I}\otimes\mathbf{X}')\mathbf{Z}\boldsymbol{\delta}]$$

ヤコービアンを無視したうえでこれを最大化してつぎの推定量をえる。

(12) Theil, p. 509. (5. 5)。なお，タイルは最小 2 乗法の枠組において，$\mathbf{I}\otimes\mathbf{X}'$ を左より掛けて変換するものを操作変数法とよんでいる。

$$(3.11)\quad \tilde{\boldsymbol{\delta}}=\{\mathbf{Z}'(\mathbf{I}\otimes\mathbf{X}')'(\boldsymbol{\Sigma}\otimes\mathbf{X}'\mathbf{X})^{-1}(\mathbf{I}\otimes\mathbf{X}')\mathbf{Z}\}^{-1}$$
$$\times\mathbf{Z}'(\mathbf{I}\otimes\mathbf{X}')'(\boldsymbol{\Sigma}\otimes\mathbf{X}'\mathbf{X})^{-1}(\mathbf{I}\otimes\mathbf{X}')\mathbf{y}$$
$$=\{\mathbf{Z}'[\boldsymbol{\Sigma}^{-1}\otimes\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Z}\}^{-1}\mathbf{Z}'[\boldsymbol{\Sigma}^{-1}\otimes\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y}$$
$$=\{\hat{\mathbf{Z}}'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I})\hat{\mathbf{Z}}\}^{-1}\hat{\mathbf{Z}}'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I})\mathbf{y}$$
$$\hat{\mathbf{Z}}=\mathrm{diag}(\hat{\mathbf{Z}}_1\,\hat{\mathbf{Z}}_2\cdots\hat{\mathbf{Z}}_L)$$
$$\hat{\mathbf{Z}}_j=\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j$$

$\boldsymbol{\Sigma}$ は通常未知である。これをつぎに述べる 2 段階最小 2 乗残差の平均交叉積の行列でおきかえると 3 段階最小乗推定量をえる。(13)

つぎに $\mathbf{P}'$ としてつぎの射影子 projection operator を使用しよう。(14)

$$\mathbf{P}'=\mathbf{I}\otimes\mathbf{H}'$$
$$\mathbf{H}'=\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

ここに $\mathbf{H}$ は $n\times n$ の対称のベキ等行列である。モデル (3.3) の左より $\mathbf{I}\otimes\mathbf{H}'$ を掛けてつぎのようにあらわしておく。

$$(3.12)\qquad (\mathbf{I}\otimes\mathbf{H}')\mathbf{y}=(\mathbf{I}\otimes\mathbf{H}')\mathbf{Z}\boldsymbol{\delta}+(\mathbf{I}\otimes\mathbf{H}')\boldsymbol{\epsilon}$$

この攪乱項はつぎの分布をもつ。

$$(\mathbf{I}\otimes\mathbf{H}')\boldsymbol{\epsilon}\sim N(\mathbf{0},\boldsymbol{\Sigma}\otimes\mathbf{H})$$

これの結合密度関数に変換 (3.12) を加えてつぎの尤度関数をえる。

$$(3.13)\quad l(\boldsymbol{\delta},\boldsymbol{\Sigma}\otimes\mathbf{H})=|(\mathbf{I}-\mathbf{C})^2|^{n/2}(2\pi)^{-Ln/2}|(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}[(\mathbf{I}\otimes\mathbf{H}')\mathbf{y}-(\mathbf{I}\otimes\mathbf{H}')\mathbf{Z}\boldsymbol{\delta}]'(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}$$
$$[(\mathbf{I}\otimes\mathbf{H}')\mathbf{y}-(\mathbf{I}\otimes\mathbf{H}')\mathbf{Z}\boldsymbol{\delta}]$$

ヤコービアンを無視して最大化することにより (3.11) と同じ結果がえられる。

$\mathbf{H}'$ が射影子であることを考慮してつぎのようにおく。

$$\hat{\mathbf{y}}=(\mathbf{I}\otimes\mathbf{H}')\mathbf{y},\quad \hat{\mathbf{Z}}=(\mathbf{I}\otimes\mathbf{H}')\mathbf{Z},\quad \hat{\boldsymbol{\epsilon}}=(\mathbf{I}\otimes\mathbf{H}')\boldsymbol{\epsilon}$$

これを用いると，いま述べたモデル，尤度関数はつぎのようにあら

(13)　Theil, p. 510 (5. 9)。

(14)　Dhrymes, pp. 209–211 は $\mathbf{P}_i'$ として $\mathbf{R}^{-1}\mathbf{X}'$ ($\mathbf{R}$ は $\mathbf{X}'\mathbf{X}=\mathbf{R}\mathbf{R}'$ を満足するもの) を用いている。

わすことができる。

(3.14)　$\hat{\mathbf{y}}=\hat{\mathbf{Z}}\boldsymbol{\delta}+\hat{\boldsymbol{\epsilon}}$

(3.15)　$$l(\boldsymbol{\delta},\boldsymbol{\Sigma}\otimes\mathbf{H})=|(\mathbf{I}-\mathbf{C})^2|^{n/2}(2\pi)^{-Ln/2}|(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}(\hat{\mathbf{y}}-\hat{\mathbf{Z}}\boldsymbol{\delta})'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}(\hat{\mathbf{y}}-\hat{\mathbf{Z}}\boldsymbol{\delta})$$

ヤコービアンを無視してこれを最大化してつぎのものをえる。

$$\tilde{\boldsymbol{\delta}}=\{\hat{\mathbf{Z}}'(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}\hat{\mathbf{Z}}\}^{-1}\hat{\mathbf{Z}}'(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}\hat{\mathbf{y}}$$

(3.16)　$$=\{\hat{\mathbf{Z}}'(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}\hat{\mathbf{Z}}\}^{-1}\hat{\mathbf{Z}}'(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}\mathbf{y}$$

この最後の式においてはつぎの関係を用いた。(15)

$$(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}\hat{\mathbf{y}}=(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}\mathbf{y}$$

ここにえた結果をみると3段階最小2乗法のもう1つの定式化に気づく。この節のはじめにおいて結合従属変数を被説明変数と説明変数とに分けたが，推定上の取扱いにあたってそれを差別することはしなかった。ここでは被説明変数には観測値をそのまま使用し，説明変数にはその推定値を用いることにしよう。こうするとつぎのモデルをえる。

(3.17)　$$\hat{\mathbf{y}}=\mathbf{Z}\boldsymbol{\delta}+\boldsymbol{\epsilon}_*$$

この新しい攪乱項 $\boldsymbol{\epsilon}_*$ がまえの $\boldsymbol{\epsilon}$ と同じ性質をもつと仮定すると，つぎの尤度関数をえる。

(3.18)　$$l(\boldsymbol{\delta},\boldsymbol{\Sigma}_*\otimes\mathbf{I})=(2\pi)^{-Ln/2}|(\boldsymbol{\Sigma}_*\otimes\mathbf{I})^{-1}|^{1/2}$$
$$\exp-\frac{1}{2}(\mathbf{y}-\hat{\mathbf{Z}}\boldsymbol{\delta})'(\boldsymbol{\Sigma}_*\otimes\mathbf{I})^{-1}(\mathbf{y}-\hat{\mathbf{Z}}\boldsymbol{\delta})$$

ここに $\boldsymbol{\Sigma}_*$ は $\boldsymbol{\epsilon}_*$ の共分散行列をあらわす。ここでは説明変数としての結合従属変数と被説明変数としての結合従属変数とを取扱上差別しているために，変換のヤコービアンが1であることは明らかである。$\boldsymbol{\Sigma}_*$ 所与のもとでのこれの最大化によって $\boldsymbol{\delta}$ の推定量として (3.16) と同じもの——$\boldsymbol{\Sigma}$ を $\boldsymbol{\Sigma}_*$ に代えている——をえる。そして $\boldsymbol{\Sigma}_*$ の推定

---

(15)　$$(\boldsymbol{\Sigma}\otimes\mathbf{H})^{-1}\hat{\mathbf{y}}=(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{H}^{-1})(\mathbf{I}\otimes\mathbf{H}')\mathbf{y}$$
$$=(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{H}^{-1}\mathbf{H}')\mathbf{y}$$
$$\mathbf{H}^{-1}\mathbf{H}'=\mathbf{H}^{-1}\mathbf{H}=\mathbf{H}^{-1}\mathbf{H}\mathbf{H}\mathbf{H}^{-1}=\mathbf{H}^{-1}\mathbf{H}\mathbf{H}^{-1}=\mathbf{H}^{-1}$$

値として (3.16) におけると同じものを使用すれば，それと同じ3段階最小乗推定量がえられる。

2段階最小2乗法

つぎにシステムの残余のものと切り離して1個の方程式を推定しよう。ここでは第1の方程式，

(3.19)
$$\begin{aligned}\boldsymbol{\epsilon}_1 &= \mathbf{y}_1 - \mathbf{Y}_1\mathbf{c}_1 - \mathbf{X}_1\boldsymbol{\beta}_1 \\ &= \mathbf{y}_1 - \mathbf{Z}_1\boldsymbol{\delta}_1\end{aligned}$$

に注目する。(式 (3.1) をみよ)[16]

はじめに，これの左より $\mathbf{P}_1'$ を掛けたものを用いる。このとき尤度関数は (3.9) より第1式のものをとり出したものとなる。そしてここでヤコービアンを無視して尤度関数を最大化することにより，推定量として

(3.20)
$$\tilde{\boldsymbol{\delta}}_1 = (\mathbf{P}_1'\mathbf{Z}_1)^{-1}\mathbf{P}_1'\mathbf{y}_1$$

をえる。これは (3.7) の第1式である。

つぎに $\mathbf{P}_1$ として $\mathbf{X}_1$ を用いると，最尤推定量として (3.11) の第1式をえる。

(3.21)
$$\tilde{\boldsymbol{\delta}}_1 = (\hat{\mathbf{Z}}_1'\mathbf{Z}_1)^{-1}\hat{\mathbf{Z}}_1'\mathbf{y}_1$$

これは (3.19) において $\mathbf{Z}_1$ を $\hat{\mathbf{Z}}_1$ に代えてえる普通の最小2乗推定量であり (3.19) へ適用された2段階最小2乗推定量に等しい。[17]

最後に $\mathbf{P}_1'$ として射影子 $\mathbf{H}'=\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ を用いる。このときにはつぎのものをえる。

$$\tilde{\boldsymbol{\delta}}_1 = (\hat{\mathbf{Z}}_1'\mathbf{Z}_1)^{-1}\hat{\mathbf{Z}}_1\hat{\mathbf{y}}_1$$

これがもう1つの2段階最小2乗推定量である。

参考　k クラス推定量

ここで $\mathbf{P}_1$ としてつぎのものを使用する場合に注目する。[18]

(16)　過剰識別の条件 $N_k < K_k + L_k$ は，標準の回帰モデルにおいて普通みたされている。Theil, p. 451。

(17)　Theil, p. 451 (5. 1)。

(18)　Theil, p. 508, 問題4. 4。

$$\mathbf{P}_1=[\mathbf{Y}_1-k\tilde{\mathbf{V}}_1\ \mathbf{X}_1]$$
$$\tilde{\mathbf{V}}=\mathbf{Y}_1-\tilde{\mathbf{Y}}_1=\mathbf{Y}_1-\mathbf{X}\tilde{\mathbf{\Pi}}_1$$

ここに $\tilde{\mathbf{\Pi}}_1$ は誘導型係数の最尤推定量（2.10）のうち第1の構造方程式に説明変数として用いられている結合従属変数 $\mathbf{Y}_1$ についてのものをあらわす。

このときの尤度関数，推定量の一般的形は操作変数法によるものと同じであるが，こうしてえる推定量が k クラス推定量とよばれている。そして $k=1$ のとき $\mathbf{P}_1=\hat{\mathbf{Z}}_1$ となり，2段階最小2乗推定量をえる。また $k=0$ のとき $\mathbf{P}_1=\mathbf{Z}_1$ となり普通の最小2乗推定量をえる。また k として適当な値を選ぶ[19]ことによって制限情報最尤推定量をえることを示すことができる。

うえで1つの方程式について述べたことをすべての方程式に拡張し，

$$\mathbf{P}'=\mathrm{diag}(\mathbf{P}_1'\,\mathbf{P}_2'\cdots\mathbf{P}_L')$$
$$\mathbf{P}_j=[\mathbf{Y}_j-k\tilde{\mathbf{V}}_j\ \mathbf{X}_j]$$
$$\tilde{\mathbf{V}}_j=\mathbf{Y}_j-\tilde{\mathbf{Y}}_j=\mathbf{Y}_j-\mathbf{X}\tilde{\mathbf{\Pi}}_j,\quad j=1,2,\cdots,L$$

を用いて変換するとシステム k クラス推定量をえる[20]。

第4節　連立方程式モデル　縮約尤度関数によるもの

最小分散法

はじめに多数方程式モデルに注目する。（1.17）をつぎのようにあらわしておく。

$$\mathbf{W}=[\mathbf{Y}\ \mathbf{X}]\quad \mathbf{A}^*=\begin{bmatrix}\mathbf{I}\\ -\mathbf{B}^*\end{bmatrix}\tag{4.1}$$

$$\mathbf{W}\mathbf{A}^*=\mathbf{E}\tag{4.2}$$

このとき尤度関数（1.18）はつぎのようにあらわされる。

$$l(\mathbf{B}^*,\mathbf{\Sigma})=(2\pi)^{-Ln/2}|\mathbf{\Sigma}^{-1}|^{n/2}\tag{4.3}$$

---

(19)　あとで述べる式（4.20）の最小根。

(20)　Savin (1973)。

$$\exp -\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}^{-1}\mathbf{A}^{*}\mathbf{W}'\mathbf{W}\mathbf{A}^{*}$$

すでに明らかなように，共分散行列の推定値は係数と同時（あるいはそのあと）でなければえることはできない。係数の推定前にえるそれの推定量は係数の関数としてのものである。

$$\tilde{\boldsymbol{\Sigma}}(\mathbf{A}^{*})=\frac{1}{n}\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^{*} \tag{4.4}$$

これを用いて縮約すると尤度関数はつぎのようになる。

$$l(\mathbf{B}^{*})=(2\pi)^{-Ln/2}|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^{*}|^{-n/2}(n)^{n/2}\exp -\frac{n^2}{2} \tag{4.5}$$

これの最大化はつぎの行列式——一般化分散——の最小化に等しい。

$$|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^{*}| \tag{4.6}$$

この方法を最小分散法とよんでおこう。ここで変換 (4.2) のヤコービアンが **1** であることに注目しておこう。

なお，参照の便のために **WA*** および **WA** をくわしく述べておこう。

$$\mathbf{W}\mathbf{A}^{*}=[\mathbf{y}_1-\mathbf{X}\boldsymbol{\beta}_1{}^{*}\ \ \mathbf{y}_2-\mathbf{X}\boldsymbol{\beta}_2{}^{*}\cdots\mathbf{y}_L-\mathbf{X}\boldsymbol{\beta}_L{}^{*}]$$

$$\mathbf{W}\mathbf{A}=[\mathbf{y}_1-\mathbf{X}_1\boldsymbol{\beta}_1\ \ \mathbf{y}_2-\mathbf{X}_2\boldsymbol{\beta}_2\cdots\mathbf{y}_L-\mathbf{X}_L\boldsymbol{\beta}_L]$$

ここでアプリオリ情報を考慮後の一般化分散の最小化による係数の推定量をもとめておこう。

$$\begin{aligned}
&\frac{1}{n}|\mathbf{A}'\mathbf{W}'\mathbf{W}\mathbf{A}|\\
&=\frac{1}{n}|[\mathbf{y}_1-\mathbf{X}_1\boldsymbol{\beta}_1\cdots\mathbf{y}_L-\mathbf{X}_L\boldsymbol{\beta}_L]'[\mathbf{y}_1-\mathbf{X}_1\boldsymbol{\beta}_1\cdots\mathbf{y}_L-\mathbf{X}_L\boldsymbol{\beta}_L]|\\
&=\frac{1}{n}\begin{vmatrix}(\mathbf{y}_1-\mathbf{X}_1\boldsymbol{\beta}_1)'(\mathbf{y}_1-\mathbf{X}_1\boldsymbol{\beta}_1) & \cdots & (\mathbf{y}_1-\mathbf{X}_1\boldsymbol{\beta}_1)'(\mathbf{y}_L-\mathbf{X}_L\boldsymbol{\beta}_L)\\ \vdots & & \vdots\\ (\mathbf{y}_L-\mathbf{X}_L\boldsymbol{\beta}_L)'(\mathbf{y}_1-\mathbf{X}_1\boldsymbol{\beta}_1) & \cdots & (\mathbf{y}_L-\mathbf{X}_L\boldsymbol{\beta}_L)'(\mathbf{y}_L-\mathbf{X}_L\boldsymbol{\beta}_L)\end{vmatrix}\\
&=|[\sigma_{gh}]|=|\boldsymbol{\Sigma}|
\end{aligned}$$

ここに σ_{gh} は $\boldsymbol{\Sigma}$ の g, h 要素をあらわす。$|\boldsymbol{\Sigma}|$ を $\boldsymbol{\beta}_i$ に関して微分し，0 に等置してつぎのものをえる。

$$\begin{aligned}
\frac{\partial|\boldsymbol{\Sigma}|}{\partial\boldsymbol{\beta}_i}&=\sum_g\sum_h\boldsymbol{\Sigma}_{gh}\frac{\partial\sigma_{gh}}{\partial\boldsymbol{\beta}_i}\\
&=2\sum_h\boldsymbol{\Sigma}_{gh}(-\mathbf{X}_i{}'\mathbf{e}_g)\\
&=2\sum\boldsymbol{\Sigma}_{gh}[-\mathbf{X}_i{}'(\mathbf{y}_g-\mathbf{X}_g\tilde{\boldsymbol{\beta}}_g)]=\mathbf{0}
\end{aligned} \tag{4.7}$$

(4.8)　$$\mathbf{e}_g=\mathbf{y}_g-\mathbf{X}_g\tilde{\boldsymbol{\beta}}_g$$

ここでつぎの関係を用いた。

$$\frac{\partial\sigma_{gh}}{\partial\boldsymbol{\beta}_i}=\begin{cases}0 & i\neq g,\ i\neq h\\ -\mathbf{X}_i'\boldsymbol{\epsilon}_h & i=g\\ -\boldsymbol{\epsilon}_g'\mathbf{X}_i & i=h,\ i=1,2,\cdots,L\end{cases}$$

そして $\boldsymbol{\beta}_i$ の解を $\tilde{\boldsymbol{\beta}}_i$ と記しておいた。$\boldsymbol{\Sigma}_{gh}$ は行列 $\boldsymbol{\Sigma}$ の g,h 要素の余因子をあらわし，$\boldsymbol{\Sigma}_{gh}/|\boldsymbol{\Sigma}|=\sigma^{gh}$（すなわち $\boldsymbol{\Sigma}^{-1}$ の g,h 要素）であるから，つぎの連立方程式をえる。

$$\begin{bmatrix}\mathbf{X}_1\cdots & \mathbf{0}\\ \vdots & \vdots\\ \mathbf{0} & \cdots\mathbf{X}_L\end{bmatrix}'\begin{bmatrix}\tilde{\sigma}^{11}\mathbf{I} & \cdots\tilde{\sigma}^{1L}\mathbf{I}\\ \vdots & \vdots\\ \tilde{\sigma}^{L1}\mathbf{I} & \cdots\tilde{\sigma}^{LL}\mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{y}_1-\mathbf{X}_1\tilde{\boldsymbol{\beta}}_1\\ \vdots\\ \mathbf{y}_L-\mathbf{X}_L\tilde{\boldsymbol{\beta}}_L\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ \vdots\\ \mathbf{0}\end{bmatrix}$$

$$\overline{\mathbf{X}}'(\tilde{\boldsymbol{\Sigma}}\otimes\mathbf{I})^{-1}\mathbf{y}-\overline{\mathbf{X}}'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}\overline{\mathbf{X}}\tilde{\boldsymbol{\beta}}=\mathbf{0}$$

これを解いて

(4.9)　$$\tilde{\boldsymbol{\beta}}=[\overline{\mathbf{X}}'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}\overline{\mathbf{X}}]^{-1}\overline{\mathbf{X}}'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}\mathbf{y}$$

をえる。これは (1.8) でえたものに等しい。

アプリオリ情報を考慮前のものについても同様の結果をえる。こうして第1節の尤度関数（の指数部分）の最大化（1.21）に等価なものとしてつぎの結果をえる。

$$\begin{aligned}&\min[\mathbf{y}-(\mathbf{I}\otimes\mathbf{X})\boldsymbol{\beta}^*](\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}[\mathbf{y}-(\mathbf{I}\otimes\mathbf{X})\boldsymbol{\beta}^*]\\ &=\min\operatorname{tr}\boldsymbol{\Sigma}^{-1}\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*\\ &=\min|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|\end{aligned}$$

完全情報最尤法

ここで多数方程式モデルより連立方程式モデルの推定に移る。ここでもこれまでと同じ記号を使用するが $\mathbf{A}^*$ の内容をつぎのように改めておく。

$$\mathbf{A}^*=\begin{bmatrix}\mathbf{I}-\mathbf{C}^*\\ -\mathbf{B}^*\end{bmatrix}=\begin{bmatrix}\boldsymbol{\Gamma}^*\\ -\mathbf{B}^*\end{bmatrix}$$

ここに上部のブロック（ヤコービアンになる部分）が $\mathbf{I}$ より $\mathbf{I}-\mathbf{C}^*=\boldsymbol{\Gamma}^*$ に変えられている。このとき尤度関数（3.4）はつぎのようにあら

わされる。(4.3) と比較してみよ。

$$(4.10)\qquad l(\mathbf{\Gamma}^*, \mathbf{B}^*, \mathbf{\Sigma}^*) = |\mathbf{\Gamma}^{*2}|^{n/2}(2\pi)^{-Ln/2}|\mathbf{\Sigma}^{-1}|^{n/2} \exp -\frac{1}{2}\mathrm{tr}\,\mathbf{\Sigma}^{-1}\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*$$

このとき，共分散行列の，係数の関数としての推定量の形は (4.4) と同じである。これを用いて縮約した尤度関数をつぎのようにあらわしておく。

$$\begin{aligned}
& l(\mathbf{\Gamma}^*, \mathbf{B}^*) \\
(4.11)\quad & = |\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|^{-n/2}|\mathbf{\Gamma}^{*\prime}\mathbf{\Gamma}^*|^{n/2}K \\
& = |\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|^{-n/2}|\mathbf{\Gamma}^{*\prime}\mathbf{Y}'\mathbf{M}'\mathbf{M}\mathbf{Y}\mathbf{\Gamma}^*|^{n/2}|\mathbf{M}'\mathbf{M}|^{-n/2}K \\
& = |\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|^{-n/2}|\hat{\mathbf{V}}'\hat{\mathbf{V}}|^{n/2}|\mathbf{M}'\mathbf{M}|^{-n/2}K \\
& = |\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|^{-n/2}|\mathbf{\Gamma}^{*\prime}\mathbf{Y}'\mathbf{Y}\mathbf{\Gamma}^*|^{n/2}|\mathbf{Y}'\mathbf{Y}|^{-n/2}K \\
& = |\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|^{-n/2}|\mathbf{\Gamma}^{*\prime}\tilde{\mathbf{V}}'\tilde{\mathbf{V}}\mathbf{\Gamma}^*|^{n/2}|\tilde{\mathbf{V}}'\tilde{\mathbf{V}}|^{-n/2}K \\
& = |\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^* + \mathbf{\Gamma}^{*\prime}\tilde{\mathbf{V}}'\tilde{\mathbf{V}}\mathbf{\Gamma}^*|^{-n/2}|\mathbf{\Gamma}^{*\prime}\tilde{\mathbf{V}}'\tilde{\mathbf{V}}\mathbf{\Gamma}^*|^{n/2} \\
& \quad |\tilde{\mathbf{V}}'\tilde{\mathbf{V}}|^{-n/2}K \\
& \quad K = (n)^{n/2}(2\pi)^{-Ln/2}\exp -\frac{n^2}{2}
\end{aligned}$$

このうち最初の式がクープマンス等によって用いられている縮約尤度関数である。

ここでつぎの関係を用いた。(最後の式はあとで使用する)

$$|\mathbf{\Gamma}^{*\prime}\mathbf{Y}'\mathbf{Y}\mathbf{\Gamma}^*|\,|\mathbf{Y}'\mathbf{Y}|^{-1} = |\mathbf{\Gamma}^{*\prime}\mathbf{\Gamma}^*|$$

$$|\mathbf{\Gamma}^{*\prime}\tilde{\mathbf{V}}'\tilde{\mathbf{V}}\mathbf{\Gamma}^*|\,|\tilde{\mathbf{V}}'\tilde{\mathbf{V}}|^{-1} = |\mathbf{\Gamma}^{*\prime}\mathbf{\Gamma}^*|$$

$$\tilde{\mathbf{Y}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

$$\tilde{\mathbf{V}} = \mathbf{Y} - \tilde{\mathbf{Y}}$$

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

$$\tilde{\mathbf{W}} = [\tilde{\mathbf{Y}}\ \mathbf{X}]$$

$$\mathbf{W} = \tilde{\mathbf{W}} + [\tilde{\mathbf{V}}\ \mathbf{O}]$$

$$\mathbf{W}' = \tilde{\mathbf{W}}' + \begin{bmatrix} \tilde{\mathbf{V}}' \\ \mathbf{O}' \end{bmatrix}$$

$$\mathbf{W}'\mathbf{W} = \tilde{\mathbf{W}}'\tilde{\mathbf{W}} + \begin{bmatrix} \tilde{\mathbf{V}}'\tilde{\mathbf{V}} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$$

$$\mathbf{W}'\tilde{\mathbf{W}}=\tilde{\mathbf{W}}'\tilde{\mathbf{W}}$$

この縮約尤度関数の最大化はつぎの分散比の最小化に等しい。

$$|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|/|\mathbf{\Gamma}^{*\prime}\mathbf{\Gamma}^*|$$

$$|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|/|\mathbf{\Gamma}^{*\prime}\mathbf{Y}'\mathbf{Y}\mathbf{\Gamma}^*| \tag{4.12}$$

$$|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|/|\hat{\mathbf{V}}'\hat{\mathbf{V}}|$$

$$|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|/|\mathbf{\Gamma}^{*\prime}\hat{\mathbf{V}}'\hat{\mathbf{V}}\mathbf{\Gamma}^*| \tag{4.13}$$

$$|\mathbf{A}^{*\prime}\tilde{\mathbf{W}}'\tilde{\mathbf{W}}\mathbf{A}^*|/|\mathbf{\Gamma}^{*\prime}\tilde{\mathbf{V}}'\tilde{\mathbf{V}}\mathbf{\Gamma}^*| \tag{4.14}$$

さらに，これに等価なものとしてつぎの制約条件つき最小化を定式化することができる。

(4.15)

最小化	条　件
$\|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*\|$	$\|\mathbf{\Gamma}^{*\prime}\mathbf{\Gamma}^*\|=c$
$\|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*\|$	$\|\mathbf{\Gamma}^{*\prime}\mathbf{Y}'\mathbf{Y}\mathbf{\Gamma}^*\|=c$
$\|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*\|$	$\|\hat{\mathbf{V}}'\hat{\mathbf{V}}\|=c$
$\|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*\|$	$\|\mathbf{\Gamma}^{*\prime}\tilde{\mathbf{V}}'\tilde{\mathbf{V}}\mathbf{\Gamma}^*\|=c$
$\|\mathbf{A}^{*\prime}\tilde{\mathbf{W}}'\tilde{\mathbf{W}}\mathbf{A}^*\|$	$\|\mathbf{\Gamma}^{*\prime}\tilde{\mathbf{V}}'\tilde{\mathbf{V}}\mathbf{\Gamma}^*\|=c$

ここでは完全情報最尤法の2つの定式化，分散比（4.12）

$$\rho=|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|/|\mathbf{\Gamma}^{*\prime}\mathbf{Y}'\mathbf{Y}\mathbf{\Gamma}^*|$$

の最小化と，それに対応する（4.15）についてつくられるラグランジュ関数

$$L=|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|-\lambda(|\mathbf{\Gamma}^{*\prime}\mathbf{Y}'\mathbf{Y}\mathbf{\Gamma}^*|-c)$$

の最小化とをとりだし，その等価性を示しておこう。

まず $\mathbf{B}^*$ の要素による最小化に注目する。$\mathbf{B}^*$ は比においては分子に含まれているにとどまり，ラグランジュ関数においては第1項に含まれているにとどまることをみると，この要素に関する偏微分を0に等値してえる $\mathbf{B}^*$ の解がこの2つの定式化において同じであることは容易にわかる。

分散比の分母およびラグランジュ関数の第2項はつぎのものの行列式であらわされる。

$\mathbf{\Gamma}^{*\prime}\mathbf{Y}'\mathbf{Y}\mathbf{\Gamma}^*$

$$=\begin{bmatrix}\mathbf{y}_1\mathbf{r}_1^* & 0 & \cdots 0\\ 0 & \mathbf{y}_2\mathbf{r}_2^* & \cdots 0\\ \vdots & \vdots & \vdots\\ 0 & 0 & \cdots \mathbf{y}_L\mathbf{r}_L^*\end{bmatrix}'\begin{bmatrix}\mathbf{I}\\ \mathbf{I}\\ \vdots\\ \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{I}\\ \mathbf{I}\\ \vdots\\ \mathbf{I}\end{bmatrix}'\begin{bmatrix}\mathbf{y}_1\mathbf{r}_1^* & 0 & \cdots 0\\ 0 & \mathbf{y}_2\mathbf{r}_2^* & \cdots 0\\ \vdots & \vdots & \vdots\\ 0 & 0 & \cdots \mathbf{y}_L\mathbf{r}_L^*\end{bmatrix}$$

（ここに $\mathbf{I}$ は単位行列をあらわす)。これは $\boldsymbol{\beta}_g^*$ を含んでいない。他方，分散比の分子およびラグランジュ関数の第1項は $\boldsymbol{\beta}_g^*$ を含んでいる。したがって $|\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*|$ の $\boldsymbol{\beta}_g^*$ に関する最小化によってえる $\boldsymbol{\beta}_g^*$ の推定量は $\mathbf{y}_g\mathbf{r}_g^*$ の線型関数（その係数行列を $\mathbf{M}_g$ とあらわす）としてあらわされる。((4.9) において $\mathbf{y}_g$ を $\mathbf{y}_g\mathbf{r}_g^*$ とかえてみよ。）この推定量を用いて縮約してつぎのものをえる。

$\mathbf{A}^{*\prime}\mathbf{W}'\mathbf{W}\mathbf{A}^*$

$$=\begin{bmatrix}\mathbf{M}_1\mathbf{y}_1\mathbf{r}_1^* & 0 & \cdots 0\\ 0 & \mathbf{M}_2\mathbf{y}_2\mathbf{r}_2^* & \cdots 0\\ \vdots & \vdots & \vdots\\ 0 & 0 & \cdots \mathbf{M}_L\mathbf{y}_L\mathbf{r}_L^*\end{bmatrix}'\begin{bmatrix}\mathbf{I}\\ \mathbf{I}\\ \vdots\\ \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{I}\\ \mathbf{I}\\ \vdots\\ \mathbf{I}\end{bmatrix}'\begin{bmatrix}\mathbf{M}_1\mathbf{y}_1\mathbf{r}_1^* & 0 & \cdots 0\\ 0 & \mathbf{M}_2\mathbf{y}_2\mathbf{r}_2^* & \cdots 0\\ \vdots & \vdots & \vdots\\ 0 & 0 & \cdots \mathbf{M}_L\mathbf{y}_L\mathbf{r}_L^*\end{bmatrix}$$

このように縮約したあとで尤度関数を $\mathbf{r}_g^*$ に関して微分してえる

$$\mathbf{r}_g^*\frac{\partial L}{\partial \mathbf{r}_g^*}=0$$

を配列して $\rho=\lambda$ をえる。これを用いると

$$L=\lambda c$$

をえる。こうして L 最小化と λ 最小化したがって ρ 最小化との等価を示すことができた。

なお，チョウは (4.12) と (4.15) によって完全情報最尤法を定式化しているが，その2つの等価性の証明は述べていない。[(21)]

3段階最小2乗法

ここでアプリオリ情報を考慮に入れる。それで行列 $\mathbf{A}$ の肩の星印を除いておく。

うえで述べたように完全情報最尤法は複雑である。それをさけるた

(21)　Chow, p. 540。

めにいろいろの単純化を加えることができる。ここに (4.13) ——星印を除く——において分母を無視して[22]つぎのものを定式化する。

$$\text{最小化}\,|\mathbf{A}'\tilde{\mathbf{W}}'\tilde{\mathbf{W}}\mathbf{A}|$$

ここに

$$\tilde{\mathbf{W}}\mathbf{A}=[\hat{\mathbf{y}}_1-\hat{\mathbf{Z}}_1\boldsymbol{\delta}_1\ \ \hat{\mathbf{y}}_2-\mathbf{Z}_2\boldsymbol{\delta}_2\cdots\mathbf{y}_L-\mathbf{Z}_L\boldsymbol{\delta}_L]$$

をあらわす。これがチョウの定式化による3段階最小2乗法である。[23]これの最小化は——分母を無視しているところにおいては—— (4.6) において述べたように，$\mathrm{tr}\,\boldsymbol{\Sigma}^{-1}\mathbf{A}'\tilde{\mathbf{W}}'\tilde{\mathbf{W}}\mathbf{A}$ の最小化に等しい。ここに $\boldsymbol{\Sigma}$ は攪乱項 $\tilde{\mathbf{W}}\mathbf{A}$ の共分散行列をあらわす。そして後者の最小化は (3.18) の指数部分の最小化に等しい。こうして3段階最小2乗法の定式化をしてつぎのものをえることができる。

$$\begin{aligned}&\min|\mathbf{A}'\tilde{\mathbf{W}}'\tilde{\mathbf{W}}\mathbf{A}|\\&=\min\boldsymbol{\Sigma}^{-1}\mathbf{A}'\tilde{\mathbf{W}}'\tilde{\mathbf{W}}\mathbf{A}\\(4.15)\qquad&=\min(\hat{\mathbf{y}}-\hat{\mathbf{Z}}\boldsymbol{\delta})'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1}(\hat{\mathbf{y}}-\hat{\mathbf{Z}}\boldsymbol{\delta})\end{aligned}$$

3段階最小2乗法を定式化するのにつぎの攪乱項にもとづくものがある。

$$\check{\mathbf{W}}\mathbf{A}=[\mathbf{y}_1-\hat{\mathbf{Z}}_1\boldsymbol{\delta}_1\ \ \mathbf{y}_2-\hat{\mathbf{Z}}_2\boldsymbol{\delta}_2\cdots\mathbf{y}_L-\hat{\mathbf{Z}}_L\boldsymbol{\delta}_L]$$

これについて，これまでと同じように結合密度関数，尤度関数，縮約尤度関数をつくることによってつぎのものをえることができる。ここに第1の式は (3.20) の指数部分の最小化をあらわしている。

$$\begin{aligned}(4.16)\qquad&\min(\mathbf{y}-\hat{\mathbf{Z}}\boldsymbol{\delta})'(\boldsymbol{\Sigma}_*\otimes\mathbf{I})^{-1}(\mathbf{y}-\hat{\mathbf{Z}}\boldsymbol{\delta})\\&=\min\boldsymbol{\Sigma}_*^{-1}\mathbf{A}'\check{\mathbf{W}}'\check{\mathbf{W}}\mathbf{A}\\&=\min|\mathbf{A}'\check{\mathbf{W}}'\check{\mathbf{W}}\mathbf{A}|\end{aligned}$$

制限情報最尤法

一層の単純化として，システムより1つの——第1の——構造方程式

(22) ヤコービアンの無視に等しい。

(23) Chow, p. 548。

(4. 17)　　$\boldsymbol{\epsilon}_1=\mathbf{Y}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1$

をとり出す。ここにアプリオリにないことのわかっているもののパラメーターを除いて次数をさげておいた。（第2節の記号と変えておいた）なお，ここでは1つの係数の値を1とする規準化を用いていない。

いま (4. 13)，(4. 14) より第1の構造方程式のものをとり出しつぎのものをえる。

(4. 18)　最小化　$[\mathbf{Y}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]'[\mathbf{Y}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]/[\tilde{\mathbf{V}}_1\mathbf{r}_1]'[\tilde{\mathbf{V}}_1\mathbf{r}_1]$

最小化　$[\tilde{\mathbf{Y}}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]'[\tilde{\mathbf{Y}}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]/[\tilde{\mathbf{V}}_1\mathbf{r}_1]'[\tilde{\mathbf{V}}_1\mathbf{r}_1]$

これが単1方程式についての最小分散化比法である。

このそれぞれに等価なものとしてつぎの制約条件つき最小2乗法を定式化することができる。[24]

最小化	条件
$[\mathbf{Y}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]'[\mathbf{Y}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]$	$[\tilde{\mathbf{V}}_1\mathbf{r}_1]'[\tilde{\mathbf{V}}_1\mathbf{r}_1]=c$
$[\tilde{\mathbf{Y}}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]'[\tilde{\mathbf{Y}}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]$	$[\tilde{\mathbf{V}}_1\mathbf{r}_1]'[\tilde{\mathbf{V}}_1\mathbf{r}_1]=c$

単1方程式についての最小分散比法と制約条件つき最小2乗法との等価性を完全情報最尤法についてと同じようにして示すことができる。

ここで，つぎのラグンジュ関数をつくる。

$$L=[\mathbf{Y}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]'[\mathbf{Y}_1\mathbf{r}_1-\mathbf{X}_1\boldsymbol{\beta}_1]-\lambda([\tilde{\mathbf{V}}_1\mathbf{r}]'[\tilde{\mathbf{V}}_1\mathbf{r}]-c)$$

これを $\boldsymbol{\beta}_1$ に関して偏微分してつぎのものをえる。

$$\bar{\boldsymbol{\beta}}_1=(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{Y}_1\mathbf{r}_1$$

これを用いてラグランジュ関数を縮約してつぎのものをえる。

$$L=[\mathbf{Y}_1\mathbf{r}_1]'\mathbf{M}[\mathbf{Y}_1\mathbf{r}_1]-\lambda([\tilde{\mathbf{V}}_1\mathbf{r}_1]'[\tilde{\mathbf{V}}_1\mathbf{r}_1]-c)$$

これを $\mathbf{r}_1$ に関して偏微分してつぎのものをえる。

(4. 19)　　$\partial L/\partial\mathbf{r}_1=2(\mathbf{Y}_1'\mathbf{M}\mathbf{Y}_1\mathbf{r}_1-\lambda\tilde{\mathbf{V}}_1'\tilde{\mathbf{V}}_1\mathbf{r}_1)=\mathbf{0}$

これの左より $\mathbf{r}_1'$ を掛けると，

$$[\mathbf{Y}_1\mathbf{r}_1]'\mathbf{M}[\mathbf{Y}_1\mathbf{r}_1]=\lambda[\tilde{\mathbf{V}}_1\mathbf{r}_1]'[\tilde{\mathbf{V}}_1\mathbf{r}_1]$$

---

(24)　チョウはさらに (4. 12) より第1の方程式のものをとり出してホテリングの問題を定式化している。

をえる。この結果を用いるとLをつぎのようにあらわすことができる。

$$L=\lambda c$$

それで，われわれの課題は（4.19）を満足する，あるいはつぎの行列式方程式を満足する最小のλをもとめることになる。

$$|\mathbf{Y}_1\mathbf{M}\mathbf{Y}_1-\lambda\tilde{\mathbf{V}}_1'\tilde{\mathbf{V}}_1|=0$$

これは制限情報最尤法について用いられているものと同じである。

2段階最小2乗法

3段階最小2乗法のところで述べた（4.15），（4.16）より第1の構造方程式のものをとり出すとつぎのものをえる。

$$\text{最小化}\quad|[\hat{\mathbf{y}}_1-\hat{\mathbf{Z}}_1\boldsymbol{\delta}_1]'[\hat{\mathbf{y}}_1-\hat{\mathbf{Z}}_1\boldsymbol{\delta}_1]|$$

$$\text{最小化}\quad|[\mathbf{y}_1-\hat{\mathbf{Z}}_1\boldsymbol{\delta}_1]'[\mathbf{y}_1-\hat{\mathbf{Z}}_1\boldsymbol{\delta}_1]|$$

これが，2段階最小2乗法の2つの定式化にほかならない。チョウはこの第1の定式化を用いている。(25)

お わ り に

この稿を閉じるにあたって，チョウの業績を吟味しておこう。

チョウはその論文「連立方程式の種々の推定量の比較」(1964) において「連立線型方程式のパラメーターを推定するために最小2乗法の自然な一般化をする。完全情報最尤法がこの一般化と同じものであることを示す。他の推定量がこの一般化からの逸脱である程度について述べる……」ことを課題としている。

チョウは普通の最小2乗法についてつぎのように述べている。(26)

「線型回帰の係数の列ベクトル y' を推定するのに，平方和，$(y-zr')'(y-zr')\equiv\|y-zr'\|^2$ すなわち，ベクトル $y-zr'$ の長さの2乗を r' に関して最小化する」p. 533。

これよりはじめてチョウは2つの方向に拡充している。その1つの

---

(25)　Chow, p. 548。

(26)　ここで引用にあたってチョウの用いた記号，式番号をそのまま使用する。

方向，便宜上yの係数の一般化をよぶものを3段階に分けて述べておこう。

スカラーβの導入

「対称式 $y\beta'-zr'$ について考える。ここに対称性はベクトルyに掛けるスカラー β' によって導入されている。……

いま述べた最小2乗法は $y\beta'-zr'$ の長さの2乗を，規準化$\beta'=1$の条件のもとに，すべての係数 β', γ' に関して最小化している。……」p. 533。

ベクトルβへの一般化

「G個の変数があるとし，それの第t観測を $y_{t1}, \cdots, y_{tG}$ と記す。($t=1, 2, \cdots, T$) これらの変数の線型結合が，固定変数 $z_{t1}, \cdots, z_{tK}$ の線型関数プラス平均ゼロ，分散 σ^2 の母集団より(z および $u_\tau, \tau \neq t$ と)独立にとられたランダム項によって生み出されていると仮定する。

$$\beta_1 y_{t1}+\cdots+\beta_G y_{tG}=r_1 z_{t1}+\cdots+r_K z_{tK}+u_t \quad t=1, \cdots, K$$

」p. 534。

「われわれの問題に適用される最小2乗法の代数は，ホテリングの問題のそれに似ている。つぎのように述べられる。

(2.2.1)　　最小化　$(\beta Y'-rz')(Y\beta'-zr')$

　　　　　条件　　$\beta Y'Y\beta'=c$　　　」p. 536。

(4.12) より1個の（第1の）方程式に関するものをとり出し，(4.19) にならって制約条件つき最小化問題を定式化するとこれをえることができる。

行列Bへの一般化

「2.2 節のモデルをG個の方程式へ一般化しよう。方程式iはつぎの通りである。

$$\beta_{i1} y_{t1}+\cdots+\beta_{iG} y_{tG}=r_{i1} z_{t1}+\cdots+r_{iK} z_{tK}+u_{ti}$$
$$(t=1, \cdots T, i=1, \cdots, G)$$

あるいは……

$$YB'=Z\Gamma'+U$$

」p. 538。

「最小2乗法のうえの一般化は制約条件つき最小化の形……で述べることができる……行列式

$$|(YB'-Z\Gamma')'(YB'-Z\Gamma')|$$

を条件 $|BY'YB|=c$ のもとに最小化する」p. 540。

これより1個の構造方程式に関するものをとり出して制限情報最尤法を定式化する。

「制限情報最尤法はつぎのことをする。

最小化　$(\beta Y_{\Delta}'-rz_1')(Y_{\Delta}\beta'-z_1r')$

条件　$\beta V_{\Delta}^{*\prime}V_{\Delta}^{*}\beta'=c$ 」p. 543。

これらの y 係数の一般化はすべて制約条件つき最小2乗法として述べられている。同じ一般化を，最小分散比法として述べている。これがチョウによる第2の方向への拡充である。

「最小2乗法を上述のように制約条件つき最小化によって定式化するほかに，説明さるべき全分散に対する残差分散の比 $\|y\beta'-zr'\|^2/\|y\beta'\|^2$ の最小化を選ぶことができる」p. 534。

「このモデルの係数を最小2乗法によって定式化するには……比 $\|Y\beta'-zr'\|^2/\|Y\beta'\|^2$ の最小化によってするであろう」pp. 534–535。

「最小2乗法のうえの一般化は……比の形においては

$$|(YB'-Z\Gamma')(YB'-Z\Gamma')|/|BY'YB|$$

を最小化する」p. 540。

チョウはこれが完全情報最尤法に等価のものであること，およびうえの条件つき最小2乗法に等価であることを証明なしに述べているにとどまる。

これより単1方程式についてのものをとり出して制限情報最尤法を定式化することができる。

「比

$$l(\beta)=\frac{\min\|Y_{\Delta}\beta'-z'r'\|^2}{\min\|Y_{\Delta}\beta'-zr_1'\|^2}$$

を β に関して最小化する。」

3段階最小2乗法，2段階最小2乗法についてはつぎのように述べている。

「この方法（3段階最小2乗法）をG個の方程式の残差の共分散行列の行列式を最小化することによって導出できることを示そう。……

このように，われわれは$G\times G$の行列 Σ^* の行列式

$$\begin{aligned}|\Sigma^*| &= |U^{*\prime}U^*| \\ &= |(u_1^*\cdots u_G^*)'(u_1^*\cdots u_G^*)|\end{aligned}$$

を $(\bar{\beta}_i, \bar{\gamma}_i), i=1, \cdots, G$ に関して最小化する」p. 548。

そして2段階最小2乗法については，その名の節において何の展開もせず，別のところでつぎのように簡単に述べているにとどまる。

「たとえば，2段階最小2乗法によって各ベクトル u_1^* の個別の長さの2乗を最小化する……」p. 548。

チョウの展開についてはつぎの批判を加えることができる。

1. 最小分散比と制約条件つき最小2乗法との等価性を制限情報最尤法において示しているにとどまっている。より一般的な完全情報最尤法についてはその等価性を示していない。

2. 3段階最小2乗法，2段階最小2乗法については，それを列挙しているにとどまり，チョウの2つの方向の一般化，分散比最小化，制約条件つき最小化のいずれからも導出されていない。

文　　献

（1） Chow, G. C. (1964) "A Comparision of Alternative Estimators for Simultaneous Equations." *Econometrica*, 32, pp. 532–553.

（2） Christ, C. F. (1966) *Econometric Models and Methods.*

（3） Dhrymes, P. J. (1970) *Econometrics. Statistical Foundations and Apllications.*

（4） Goldberger, A. S. (1964) *Econometric Theory.*

（5） Johnston J. (1960) *Econometric Methods.*

（6） Kmenta, Jan and Roy Gillbert (1968) "Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions." *Econometrica.*

（7） Koopmans, T. C., Wm. C. Hood, (1954) "The Estimation of Simultaneous Linear Economic Relations," Wm. C. Hood and T. C. Koopmans, (editor) *Studiss in Econometric Methods. Cowles Commission for Research in Econometrics. Monograph 14.*

(8) Koopmans, T. C., H. Rubin and R. B. Leipnik (1950) "Measuring the Equation Systems of Dynamic Economics." Chapter 11 of Koopmans, T. C. (editor) *Statistical Inference in Dynamic Economic Models.*

(9) Malinvaud, E. (1966) *Statistical Methods of Econometrics.*

(10) Savin, N. E. (1973) System *k*-Class Estimators, *Econometrica* 41 pp. 1125–1136.

(11) Theil, Henri (1971) *Principles of Econometrics.*