

観測誤差に関する若干の問題

地主重美

筆者に別の機会に⁽⁵⁾、変数に観測誤差に含んでいる構造方程式で、パラメータ推定に関する古典方法——観測誤差項の確率分布に関して特殊の強い仮定がなされる方法——をさらに拡張し、ややゆるめた仮定のもとで推定できる可能性についてのべた。しかし、経済学の最も興味ある主題が予測の問題であることから、このペーパーでは第一に上の拡張された方法を予測問題に応用してみる。第二に多変数の関係式に応用した場合の例についてのべ、最後に上のペーパーではふれなかった他の二つの推定方法についても若干のべてみたい。

1. 予測問題への拡張

さきのペーパーでは、“真”の変数値の間の関係式におけるパラメータの推定に重点をおいて議論を展開してきたが、ここでは所与のXに対してYを予測すること、同じことであるが所与のXに対してYの真値 Ψ を予測することが当面の問題である。たとえばYを消費支出の観測値、Xを所得の観測値とし、消費支出と所得の真値の間には一期間のラグをもつ次のような関係式が仮定されるものとしよう。

$$\Psi_t = \alpha + \beta X_{t-1}$$

たとえば第 n 期についてみると、所得の観測値 X_n は既知であるがその真値 X_n は未知であり、しかもわれわれは Y_{n+1} (又は Ψ_{n+1})を予測しなければならない。さきのペーパーでのべた拡張された古典的方法によって α 、 β を推定し、これを用いて予測

$$(1) \hat{Y}_{n+1} = \hat{\alpha} + \hat{\beta}X_n$$

を行なうべきであろうか。あるいはX、Yの観測値に直接最小自乗法を施し、この最小自乗関係式から予測を行なうべきだろうか。ここでわれわれが考えて

いるモデルは次のようなものである。

$$(2) X = X + u$$

$$(3) Y = \Psi + v$$

$$(4) \Psi = \alpha + \beta X + \varepsilon$$

したがって

$$(5) Y = \alpha + \beta X + v + \varepsilon$$

である。それ故予測期望値は明らかに、

$$(6) E(Y_{n+1} | X_n) = E(\alpha + \beta X_n + \varepsilon_n + v_n | X_n) = \alpha + \beta E(X_n | X_n)$$

したがって $E(X_n | X_n)$ を求めなければならない。いま、(2)式において、 X 、 u はそれぞれ平均値 \bar{X}_n が零、分散 σ_x^2 、 σ_u^2 をもち、独立の正規分布をしていると仮定しよう。

$$\sigma_x^2 = \sigma_x^2 + \sigma_u^2$$

を考慮してわれわれは、

$$E(X_n | X_n) = \frac{\sigma_u^2 \bar{X}_n + \sigma_x^2 X_n}{\sigma_x^2 + \sigma_u^2}$$

をうる。(6)式は

$$(7) E(Y_{n+1} | X_n) = \alpha + \frac{\beta \sigma_u^2 \bar{X}_n}{\sigma_x^2 + \sigma_u^2} + \frac{\beta \sigma_x^2 X_n}{\sigma_x^2 + \sigma_u^2}$$

(1)式と(7)式の比較から直ちに明らかなように、(1)式は期望値から解離してをりさきにのべた「拡張された古典的方法」は適切ではない。

n 個の標本観測値に直接最小自乗法を適用して β を推定すると

$$b_n = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$= \frac{\Sigma(X - \bar{X})(\Psi - \bar{\Psi}) + \Sigma(X - \bar{X})(v - \bar{v}) + \Sigma(\Psi - \bar{\Psi})(u - \bar{u}) + \Sigma(u - \bar{u})(v - \bar{v})}{\Sigma(X - \bar{X})^2 + 2\Sigma(X - \bar{X})(u - \bar{u}) + \Sigma(u - \bar{u})^2}$$

ここで誤差が相互に独立であり、それぞれの真の変数値からも独立であると仮定される場合には、分子の最後の三項、分母の第二項は標本が大きくなるにつれて零に近づくから、

$$\begin{aligned}\text{Plim } b_n &= \frac{\beta \Sigma (X - \bar{X})^2}{\Sigma (X - \bar{X})^2 + \Sigma (u - \bar{u})^2} \\ &= \frac{\beta}{1 + \sigma_u^2 / \sigma_X^2}\end{aligned}$$

いま最小自乗予測値は

$$(8) \quad \hat{Y}_{n+1} = \bar{Y} + b(X_n - \bar{X})$$

として示されるから、この予測値の確率極限は、

$$\begin{aligned}(9) \quad \text{Plim } \hat{Y}_{n+1} &= \alpha + \beta \bar{X}_n + \frac{\beta}{1 + \sigma_u^2 / \sigma_X^2} (X_n - \bar{X}_n) \\ &= \alpha + \frac{\beta \sigma_u^2 \bar{X}_n}{\sigma_X^2 + \sigma_u^2} + \frac{\beta \sigma_X^2}{\sigma_X^2 + \sigma_u^2} X_n \\ &= E(Y_{n+1} | X_n)\end{aligned}$$

したがって、このタイプの予測が必要とされる場合には、最小自乗法は、構造パラメーターの推定値を求むるのに適切な方法とはいえないにしても、予測の方法としてはむしろ適当なものといつてよい。

2. 多変数問題への拡張⁽²⁾

一般化された古典的方法を二個以上の変数を含む関係式に拡張することはいわば形式的な問題である。ただ通常の方法は、単に“真”の変数の間の exact な函数関係だけを問題としているが、ここでは諸変数間の確率的関係、すなわち観測誤差をもつ諸変数間の関係式を考える。はじめに三変数を含む関係式の場合を考え、そのあとで多変数のより一般的な場合に進むことにしよう、ここで次のようなモデルを仮定する。

$$(10) \quad X_{1t} = X_{1t} + u_{1t}$$

$$(11) \quad X_{2t} = X_{2t} + u_{2t}$$

$$(12) \quad Y_t = \Psi_t + v_t$$

さらに

$$(13) \quad \Psi_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$$

ここで u_1 , u_2 , v はいずれも観測誤差を示し、 ε は確率的攪乱項である。いま u_1 , u_2 , v はそれぞれ、分散 $\text{var}(u_1)$, $\text{var}(u_2)$, $\text{var}(v)$, 共分散 $\text{cov}(u_1, u_2)$,

$cov(u, v)$, $cov(u, v)$ をもち、平均値零のまわりに正規分布をしているものと仮定する。又 ε は平均値零、分散 $var(\varepsilon)$ で正規分布をしてをり、観測誤差 u_1 , u_2 , v からは独立に分布しているものと仮定しよう。さらに四箇の誤差項はともに X_1 , X_2 の分布からは独立であると仮定する。(12), (13)から

$$\begin{aligned} (14) \quad & var(X_1) = var(X_1) + var(u_1) \\ & var(X_2) = var(X_2) + var(u_2) \\ & var(Y) = var(\Psi) + var(v) \\ & \quad = \beta_1^2 var(X_1) + \beta_2^2 var(X_2) \\ & \quad \quad + 2\beta_1\beta_2 cov(X_1, X_2) + var(\varepsilon) + var(v) \\ & cov(X_1, X_2) = cov(X_1, X_2) + cov(u_1, u_2) \\ & cov(Y, X_1) = \beta_1 var(X_1) + \beta_2 cov(X_1, X_2) + cov(u, v) \\ & cov(Y, X_2) = \beta_1 cov(X_1, X_2) + \beta_2 var(X_2) + cov(u_2, v) \end{aligned}$$

となる。もし X_1 , X_2 が二変数正規分布をとるならば、 Y , X_1 , X_2 は多変数正規分布をとることになり、第二次積率は

$$MX_1, X_2 = \frac{1}{n} \sum_{t=1}^n (X_{1t} - \bar{X}_1)(X_{2t} - \bar{X}_2)$$

は(14)式の左辺による共分散の最大推定値となる。他の共分散、分散についても同様である。もし、観測誤差の分散・共分散行列が予め知られていると仮定すれば、(14)式の第一、第二、第四方程式を第五、第六方程式に代入することによって $\hat{\beta}_1$, $\hat{\beta}_2$ に関して項の同時方程式がえられる。

$$\begin{aligned} (15) \quad & \hat{\beta}_1 [MX_1, X_1 - var(u_1)] + \hat{\beta}_2 [MX_1, X_2 - cov(u, u_2)] \\ & \quad = MYX_1 - cov(u, v) \\ & \hat{\beta}_1 [MX_1, X_2 - cov(u, u_2)] + \hat{\beta}_2 [MX_2, X_2 - var(u_2)] \\ & \quad = MYX_2 - cov(u_2, v) \end{aligned}$$

ここで(15)式は一次正規の場合の最小自乗推定式

$$(X'X)b = X'Y$$

にきわめて類似している。この類似性は二変数の場合を例にとると一層明らかになる。いま二変数の偏差を考えて

$$b_1 MX_1, X_2 + b_2 MX_1, X_2 = MX_1, v$$

$$b_1 MX_1, X_2 + b_2 MX_2, X_2 = MX_2, v$$

がえられる。上式とさきの(15)式の間の唯一の相違は、(15)において標本二次積率が、誤差項の分散・共分散項を用いてかき改められていることである。

さて一般的ケースに進もう。いま説明変数 X 、被説明変数 Y を偏差で表わすと

$$X \equiv \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{pmatrix}$$

$$Y \equiv \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad V \equiv \begin{pmatrix} cov(u_1, v) \\ cov(u_2, v) \\ \vdots \\ cov(u_k, v) \end{pmatrix}$$

$$U \equiv \begin{pmatrix} var(u_1) & cov(u_1, u_2) & \cdots & cov(u_1, u_k) \\ cov(u_1, u_2) & var(u_2) & \cdots & cov(u_2, u_k) \\ \vdots & \vdots & \ddots & \vdots \\ cov(u_1, u_k) & cov(u_2, u_k) & \cdots & var(u_k) \end{pmatrix}$$

したがって

$$\hat{\beta} = \{\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k\}$$

に対する一般推定方程式は

$$(16) \quad (X'X - U) \hat{\beta} = X'Y - V$$

切片項 β_0 は標本平均を計算して求められることはいうまでもない。確率的攪乱項の分散の推定値は、(14)の第三式を拡張して次のような一般式で表わすことができる。

$$(17) \quad \hat{\sigma}_\varepsilon^2 = M_{Y,Y} - vav(v) - \hat{\beta}'(X'X - U)\beta$$

かくて観測誤差を含む多変数の関係式において、 $\hat{\beta}$ を推定する一般式は(16)であり、確率攪乱項 ε を推定する一般式は(17)である。われわれは古典的最小自乗法を若干拡張することによって、観測誤差を含む関係式の回帰係数の推定について最小自乗法を直接に適用する場合よりもより有効な推定値をうることができるのである。

4. 他の二つの推定方法

観測誤差を含む変数による関係式（ここでは一次正規の回帰モデルを考えているのであるが）のパラメーターを推定する古典的方法については別の機会に詳しくのべたから、ここではそれと違った一種の簡便法を説明する。

その第一は観測値をグループ分けしてパラメーターの推定を行なう方法で、ワルド、パートレットによって提唱され、誤差の分散・共分散行列に関してなんら特定の仮定をおくことなしに一次の構造方程式のパラメーターについて一貫性をもつが、有効性をもたない推定を可能にする。まずワルドの方法からのべてみよう⁽⁴⁾。次のモデルを仮定する。

$$X = X + u$$

$$Y = \Psi + v$$

$$\Psi = \alpha + \beta X$$

ここで u, v は相互に、かつまた系列的に独立であると仮定しよう。観測数 n が偶数すなわち $n=2m$ と仮定すれば、実際の手順はすこぶる簡単になる。まず変数 X をその数値の大いさの順序に並べてみる。 X の下ヅキがその順序を表わすものとすれば、次のような順序づけられた X の数列がえられる。

$$X_1, X_2, X_3, \dots, X_m, X_{m+1}, \dots, X_n$$

X の対応する値にしたがって順序づけられた Y の数列は

$$Y_1, Y_2, Y_3, \dots, Y_m, Y_{m+1}, \dots, Y_n$$

である。いま X の数列を X_1 から X_m までと、 X_{m+1} から X_n までの二つのグループに分割し、それぞれの X に対応する Y も二つのグループに分たれる。このサブ・グループの平均値を、

$$\bar{X}_1 = \frac{1}{m} \sum_{i=1}^m X_i$$

$$\bar{X}_2 = \frac{1}{m} \sum_{i=m+1}^n X_i$$

$$\bar{Y}_1 = \frac{1}{m} \sum_{i=1}^m Y_i$$

$$\bar{Y}_2 = \frac{1}{m} \sum_{i=m+1}^n Y_i$$

と定義する。ワルドのサジェストしている推定値は

$$(18) \quad \hat{\beta} = b = (\bar{Y} - \bar{Y}_2) / (\bar{X}_1 - \bar{X}_2)$$

$$(19) \quad \hat{\alpha} = a = \bar{Y} - b\bar{X}$$

である。この推定値が一致性をもつことを示すために次式の ($n \rightarrow \infty$ のときの) 下極限が正であると仮定する。

$$(20) \quad \liminf_{n \rightarrow \infty} \left| \frac{1}{m} \sum_{i=1}^m X_i - \frac{1}{m} \sum_{i=m+1}^n X_i \right| > 0$$

ワルトはさらに、 α, β に対する信頼区間がいかにきめられたらよいかを示している。しかし、不幸にして条件(20)は正規分布をなす変数の場合には妥当しない。つまり変数 X_i がランダムにとられるならば、条件(20)は成立しないだろう。ワルトの推定方法は、誤差の分散・共分散行列に関してなんらの仮定なしに未知のパラメーターの推定を可能にするが、条件(20)が強すぎるために一般的に応用できる手法とはいいがたい。

そこでパートレット⁽¹⁾はこれに修正を加え、大いさの順に順序づけられた X の数列と、それに対応する Y の数列のそれぞれの最後から k 個の観測値をとりそのサブ・グループの平均をとってそれを次のように定義する。

$$\begin{aligned} \bar{X}_1 &= \frac{1}{k} \sum_{i=1}^k X_i & \bar{X}_3 &= \sum_{i=n-k+1}^n X_i \\ \bar{Y}_1 &= \frac{1}{k} \sum_{i=1}^k Y_i & \bar{Y}_3 &= \sum_{i=n-k+1}^n Y_i \end{aligned}$$

構造係数の推定値は、

$$(21) \quad b' = (\bar{Y}_3 - \bar{Y}_1) / (\bar{X}_3 - \bar{X}_1)$$

$$(22) \quad a' = \bar{Y} - b' \bar{X}$$

である。ここでサブ・グループの平均を求むるに用いられる観測数を決定しなければならぬが、 X が等間隔に配列されているような特別の場合にかぎり、所与の誤差分數に対して、 b' の標本分散は極小になることをパートレットは示している。

グループ分けの方法は、誤差の分散・共分散行列に関して特別の仮定をおくことなしに、線型の構造関係式のパラメーターを推定する、一致性はもつがしかし有効性のない方法であり、古典的方法とは大きな違いがある。第一に、この方法は計算が頗る単純であり、その上正規性の仮定をおく必要がない点ですぐれているが、第二に古典的方法が多変数の場合に容易に拡張できるのに対し、

二変数以上を含む関係式にまで拡張することができないために、實際上その利用価値は大いに減殺される。

グループ分けの方法と並んで単純な推定法としてあげられるのが操作変数 (Instrumental Variables) による方法である⁽²⁾⁽³⁾。これは観測誤差のある変数の線型構造関係式で、操作変数を使用してパラメーターの推定を行なうもので、予め誤差分散・共分散行列の知識なしに一致推定値を求めるのに適用できる。いま

$$X = X + u$$

$$Y = Y + v$$

$$Y = \alpha + \beta X$$

なるモデルを仮定すれば

$$Y = \alpha + \beta X + w$$

ここで $w = v - \beta u$ である。この最後の式に最小自乗法を直接適用したのでは、最良線型不偏一致推定値をうることはできない。これを証明することに容易である。まず X と w の共分散

$$E\{w[X - E(X)]\} = E(v - \beta u)u = -\beta \text{var}(u)$$

をとると、この値が零にならないことは不偏性の欠如を示し、又

$$\text{Plim } b = \frac{\beta}{1 + \sigma_u^2 / \sigma_X^2}$$

で $\text{Plim } b \neq \beta$ であることは一致性の欠如を表わしている。もし、誤差 u, v のいずれからも独立なある変数 Z をみつけることができるならば、次のような推定値を考えることができる。

$$(23) \quad \hat{\beta} = \frac{\sum_{t=1}^n y_t z_t}{\sum_{t=1}^n x_t z_t}$$

ここで小文字 x, y, z は平値からの偏差を表している。したがって、

$$\hat{\beta} = \frac{\beta \sum_{t=1}^n x_t z_t + \sum_{t=1}^n z_t (W_t - \bar{W})}{\sum_{t=1}^n x_t z_t}$$

z は誤差 u, v のいずれからも独立であると仮定されているから、

$\sum_{t=1}^n z_t (W_t - \bar{W})$ の項は、 n が大きくなるにつれて、しだいに零に収束する。そ

れ故、(23)式は β の一致推定値を与える。この Z がここでいう操作変数に外ならない。単一方程式に含まれている未知のパラメーターの推定に操作変数を使用する方法について既に多くの学者によって試みられたところであるが⁽²⁾⁽³⁾、簡単にその手続を紹介してみよう⁽²⁾。まず与えられた方程式における内生変数の係数となっているパラメーターの数を数え、それに等しい個数の外生変数（もっとも、おくれを含んだ内生変数を含めて、先決変数まで操作変数として選ぶことも可能である）を当該方程式には含まれないが、体系内の他の外生変数から操作変数としてとり出す。この外生変数の一つをこの方程式の各変数に乘じすべての標本観測値にわたって加え合わせる。この手続を操作変数として選定された外生変数全部についてくりかえすことによって、与えられた方程式に含まれる未知パラメーターの個数と同じ個数の、未知パラメーターに関する線型方程式がえられる。この連立方程式を未知のパラメーターについて解くことによって、それらの推定値を求めるのである。操作変数による方法の論理はきわめて単純で、単一方程式

$$(24) \quad y = \beta x + w'$$

において直接に x と y との関係を考察する代りに、この方程式には含まれていない体系内の他の外生変数 Z が経済体系の運動を通じて x および y に影響するという迂回の径路を通して x と y との関係を把握しようというのである。ヴァラヴァニス教授は前者をモデル(24)の manifest part, 後者を latent part とよんでいる⁽³⁾。

(24)式で

$$\beta = \frac{\Delta y}{\Delta x} = \frac{\Delta y / \Delta z}{\Delta x / \Delta z}$$

$\frac{\Delta y}{\Delta x}$ は manifest part を, $\frac{\Delta y / \Delta z}{\Delta x / \Delta z}$ は latent part を表わしている。いま

$$\gamma_1 = \frac{M_{yz}}{M_{zz}} \quad \gamma_2 = \frac{M_{xz}}{M_{zz}}$$

と定義すれば

$$\hat{\beta} = \frac{\gamma_1}{\gamma_2} = \frac{M_{yz}/M_{zz}}{M_{xz}/M_{zz}} = \frac{M_{zy}}{M_{zx}}$$

つまり latent モデルの未知パラメーターを最小自乗法で推定する手続を通して

β を推定するというのが操作変数による方法の論理であり、そのかぎりではこの推定方法も正当化されうるように思われる。この方法は又、reduced form に最小自乗法を適用する手法と密接な関係がある。いま次のような complete model を考えてみる。

$$y = \beta x + w'$$

$$\frac{1}{r_1} y + \frac{1}{r_2} x - z = v$$

明らかに第一式は manifest part を、第二式は latent part を示している、reduced form は

$$Dy = \beta z + \frac{1}{r_2} w' + \beta v$$

$$Dx = z - \frac{1}{r_1} w' + v$$

ここでDは行列式

$$\begin{vmatrix} 1 & -\beta \\ \frac{1}{r_1} & \frac{1}{r_2} \end{vmatrix}$$

である。

しかし、操作変数による推定方法には種々なる困難があり、ここではそのうちの三つを指摘しておこう。第一は、操作変数として選定される変数の性質に恣意性があることである。推定されるパラメーターの数とちょうど同数の操作変数が、与えられた方程式の外にある場合のように、選択される操作変数の組が唯一つという特殊な場合には、与えられた方程式はちょうど識別可能であり推定方法もきわめて容易である。しかし、操作変数の組が、与えられた方程式に対して一個以上ある一般的な場合には、操作変数として選択された特定の組によって推定値も決ってしまうため、この選定の任意性がこの推定法の価値を低めることになる。そこで操作変数の選定に対してある基準を設けることが必要になる。そこで第一に、数個の操作変数を必要な場合には、その間に高度の相関をもつようなものは避けなければならない。第二に与えられた方程式の内生変数と最も高い相関をもつような変数の組を選ぶべきである。したがって又第三に当該方程式の中にすでに存在している先決変数と最も低い相関をも

つような組が選ばれるべきである。かくして選定の任意性をかなり限定し、手早い結果の期待できる操作変数による推定法の応用範囲を拡げることができる。操作変数を使用する推定方法に伴なう困難の第二は、操作変数の選定に当ってそれぞれが互いに独立であり、かつ又観測誤差からも独立であることが仮定されていたのであるが、この独立性の有無をチェックすることがきわめて困難であるということである。第三は、この方法によって推定されるパラメーターは不偏性をかくが、一致性をもっていることを特質としているにもかかわらず、きわめて大きい標本分散の可能性を含む場合には、その特質は保証されない。

参 考 文 献

- (1) M. S. Bartlett, "The Fitting of Straight Lines if Both Variables are subject to Error," *Biometrics*, 1949, pp. 207—42.
- (2) L. R. Klein, *A Textbook of Econometrics*, Part III & VII.
- (3) S. Valavanis, *Econometrics*, chap. 7.
- (4) A. Wald, "The Fitting of Straight Lines if Both Variables are subject to Error," *Annals of Mathematical Statistics*, 1940 pp. 284—300.
- (5) 地主重美, 「計量経済学における観測誤差の問題」, 季刊理論経済学 1962年5月