回帰残差の検討

野　口　和　也

1. 確率モデルの構築

経済理論で用いられる'model'という用語は，諸変数間に成立つ一連の関係を仮説的に表現したものであり，これらの関係は方程式（体系），グラフ，表，等々によって示される。ここで，仮説的表現を方程式（体系）の構築のみに限定して考えることにすると，その主体は経済現象を数学的に定式化することであって，この段階で最も重要な役割を演ずるのは，現実の経済に対する深い洞察と思考演繹である。

例として消費関数をとりあげてみよう。消費 C が所得 Y の関数であるということは

(1.1)　$C=f(Y)$

と表わされる。さらに現実の経済に対する"深い洞察"によって次のようなことが付け加えられるであろう。

(1.2)　$0<f'(Y)<1$

この式の意味するところは，「所得が増加すれば消費も増加するであろう。しかし増加率は所得のそれよりも小さい」である（f' は1階の導関数)。さらに「所得の増分に対応する消費の増分は所得の増加にともない減少する」という仮定

(1.3)　$f''(Y)<0$

が設定される。ここで注意すべきことは，現段階では未だ変数間の"定量的"な関係（例えば，一変数の一定量の変化が他の変数にどれだけの大きさの変化をおよぼすかというような）については何らの叙述もなされておらず，経済変

数間の“定性的”関係（例えば，変数間の変化の方向が等しいか否かというような）が示されているにすぎない，ということである。

これに対して，経済変数間の関数関係を“定量的”に把握しようとするためには，上述のモデル（1.1)〜(1.3）をさらに展開することが必要である。計量経済分析における‘model’は統計データが生成される構造を何らかの形で近似表現したものでなくてはならない。現実のデータに統計学の手法を応用するためには，まず第一に「関数型の想定(特定化)：specification」を行なう。消費関数の例では（1.1）を

(1.4)　$C=f(Y)=\alpha+\beta Y$　（α, β は定数）

とし，(1.2）を

(1.5)　$0<\beta<1$

とすることが方法のひとつ（線型近似）である。さらに仮定（1.3）を考慮に入れるのであれば，(1.4）と（1.5）の代りに

(1.6)　$$\begin{cases} \log C=\alpha+\beta\log Y \\ 0<\beta<1 \end{cases}$$

とすればよい。関数型の想定が経済現象にかんする理論的・先験的知識によるところが大きいことは言うまでもないが，計量分析の成否はこのようなモデルの想定によって大きく左右されるのが常である。

計量分析のために次に必要となるのは経済モデル（1.6）を確率モデル（stochastic model)(1) へ展開することである。このことは，いくつかの経済変数を確率変数と見なし，さらに方程式に誤差項（error term）をとり入れることを指す。この誤差項とは，経済モデルによる単純化のために生ずる現実とのギャップを調整するもの，と考えればよい。(1.4）を例にとり，時点 i における誤差を e_i で示せば，

(1.7)　$C_i=\alpha+\beta Y_i+e_i$

となり，e_i は「所得以外の無数の要因が（時点 i における）消費支出に及ぼす影響と線型近似による誤差」をまとめたものを表わす。確率モデルとは誤差項が確率変数であるようなモデルであり，誤差項の「分布型の想定」が必要に応

じて行なわれる。

さて，前述の関数型の想定は，現実の経済を可能な限り適切に描写するように行なわなければならないのは勿論のことであるが，さらに誤差項分布の型の想定もまた現実のデータにうまくあてはまるものであることが望ましい。とはいえ，統計的方法にも限界があるから，ある程度の単純化はやむを得ず，結局のところ，このふたつの“想定”の妥当性，すなわち確率モデルの構築の適否の判断は相対的なレベルにおいてのみ行なわれざるを得ない。かくして適否の判断そのものが分析目的や分析家の主観，等々によって異なることになり，そのこと自体現時点では解決しようのない問題となっている。この問題に対する統計理論側からの接近は，現在のところ，観測値と確率モデルとのあいだに統計的に有意な矛盾が認められるか否かによっている(2)。有意な矛盾が認められないときモデルは真 (true) であり，有意な矛盾が認められるとき偽 (false) であるといい，特に関数型の想定に誤りがある場合，モデルに“想定誤差 (specification error)”が含まれているという。本論は確率モデルの真偽判定を，誤差項分布型の想定の妥当性を中心に，最小二乗残差を用いて記述統計的に検討する方法について論じるものである。

2. 標準線型回帰モデルの仮定

計量分析において主要な役割を演じている回帰モデル（確率モデルの一種）に関して，誤差項分布の想定の妥当性を検討してみよう。次のような p 個の独立変数 x_i $(i=1, 2, \dots, p)$ を含む回帰モデルを考える。

$$(2.1)\quad \underset{(n\times1)}{\boldsymbol{y}} = \underset{(n\times p)(p\times1)}{X\boldsymbol{\beta}} + \underset{(n\times1)}{\boldsymbol{\varepsilon}}$$

n は観測値数，$\boldsymbol{y}$ は従属変数の観測値ベクトル，$\boldsymbol{\varepsilon}$ は誤差項ベクトルを示す。

一般に，モデル (2.1) に対して与えられる仮定(3) は次のようなものである。

仮定 1　モデルはパラメータ $\boldsymbol{\beta}$ にかんして線型関数である。

仮定 2　$\mathrm{rank}(X)=p$

仮定 3　$E(\boldsymbol{\varepsilon})=0$

仮定 4　$Cov(\boldsymbol{\varepsilon})=\sigma^2 I$　（σ^2 は定数）

仮定 5　$E(\varepsilon_i\varepsilon_j)=0,\ i\neq j$

仮定 6　確率変数ベクトル $\boldsymbol{\varepsilon}$ の分布は $\boldsymbol{X}$ および $\boldsymbol{\beta}$ とは無関係である。

仮定 7　$\boldsymbol{\varepsilon}\sim N(\boldsymbol{o},\ \sigma^2\boldsymbol{I}_n)$

このうち，仮定1～6を満たすモデルを「標準線型回帰モデル（standard linear regression model)」，仮定1～7を満たすモデルを「標準線型正規回帰モデル（standard linear normal regression model)」と呼び，どちらもよく用いられる。

仮定2は最小二乗推定量[4]

(2.2)　$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$

の存在のために必要な条件であり，$\mathrm{rank}(\boldsymbol{X})=\mathrm{rank}(\boldsymbol{X}'\boldsymbol{X})=p$ により $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ が非特異行列となり，モデル（2.1）には一義的な解（2.2）の存在することが証明される。この仮定が満たされない場合，独立変数間に多重共線性（multi-collinearity）が生じたという[5]。多重共線性の問題は前節で述べた問題と並んで計量経済学では重要な研究テーマであり，ここでは多重共線性に対する完全な解決策は未だ現われていない，と言うにとどめておこう。

仮定3～7はいわゆる「誤差項分布にかんする仮定」である。仮定3は誤差の値が平均的にゼロであること，すなわちプラスの値をとる可能性とマイナスの値をとる可能性が均等であることを表わしている。このことは，誤差項の他に未知の定数項がモデルの右辺に加えられていると考える（例えば $x_1\equiv 1$ とすれば β_1 が定数項となる）ことによって，$E(\boldsymbol{\varepsilon})$ がゼロベクトル以外の定数ベクトルである可能性は一般性を失うことなく排除される。仮定3の違背がおこるのは，$\boldsymbol{y}$ の変動に対して重要な意味を持つ他の独立変数 x_{p+1} を除外してしまっている場合である。この場合，真のモデルは，$\boldsymbol{y}=\boldsymbol{X\beta}+\beta_{p+1}x_{p+1}+\boldsymbol{\varepsilon}'$ であるから $\boldsymbol{\beta}$ についてのみ最小二乗推定量を求めても $E(\boldsymbol{\varepsilon}')=\boldsymbol{0}$ とならないことは明らかである。

仮定4は均一分散性（homoscedasticity）の仮定と呼ばれるもので，次の仮定5とともに最小二乗推定量の“良さ”を証明するうえで本質的な役割を果たすものである。均一分散性とは，バラツキの尺度である分散の大きさが各観測

点を通じて一定であることを意味し，また仮定５は各観測点の間に相関関係がないことを述べている。例として前節の消費関数をとりあげよう。時系列分析を考え，添字 i で時間を表わすものとすれば，単純に

(2.3)　$C_i=a+bY_i+e_i,\ i=1,\ 2,\ \dots,\ t$ $\left(\begin{array}{l}i\text{は変数の種類を表わす}\\ \text{ものでないことに注意}\end{array}\right)$

と書ける。仮定４は分散が時間とともに拡大しないことを要求し，また仮定５はある任意の時点 j における誤差項の期待値 $E(e_j)$ が，その k 期前の $j-k$ 期における誤差項の期待値 $E(e_{j-k})$ と独立であることを意味する($k=1,\ 2,\dots$)。したがって，仮定５によれば，今期の消費水準が所得によって決定される期待値を上まわった（誤差項がプラスの値をとった）としても，来期の消費水準が期待値を上まわる可能性が高くなるとは限らない。誤差項が仮定３～５を満足するような時系列を「定常時系列（stationary time series)」と呼び，逆に仮定４が満たされない場合「系列相関」（serial correlation)」が存在すると言う。

現実のデータが，このふたつの仮定を満足するものであるか否かは非常に疑わしいものであろう。例えば，所得水準の上昇は消費における“気まぐれ”的誤差要因部分を拡大するであろうし，あるいは一度引き上げられた消費の水準は再びもとの水準に低下することは容易でないかも知れない。かくして，こと経済データにかんする限り，仮定４，５は妥当でないと思われるケースが比較的多いと言えよう。

仮定６の誤差項の確率分布と独立変数との間が無関係である（したがって誤差項分布とパラメータとの間も無関係）という仮定も，経済モデルとしては非現実的な場合が多い。上の例でいえば，誤差項分布のバラツキは所得 Y の水準に依存すると考えられる。

かくして，仮定３～７を同時に満たすような観測値データは仲々現実には存在し難いものであると思われるが，とはいえ標準線型回帰モデルの想定そのものを否定してしまう態度もまた疑問である。なんとなれば，分析に用いられるデータが，これらの仮定を満足しているか否かを判断するための先験的な知識などとうてい持ち得ないし，もとより経済理論そのものも完全なものではない。

結局のところ，標準線型回帰モデルのための仮定は経済モデルを確率モデルに展開していく際の，「単純化のための仮定」であり，それによって最小二乗推定量は，Gauss-Markoff の定理によって“最良不偏線型推定量（best linear unbiased estimator : BLUE)”であることが示され，統計理論上きわめて優れた推定量となりうるのである。さらに仮定７を設定すれば，信頼区間の構成や仮設検定の準備を進めることができる。

以上のような考え方からすれば，個々の仮定の妥当性にのみ固執せず，なによりもまず「単純化のための仮定」を出発点として，単純な最小二乗推定を行なってみることが望ましい。しかる後，その結果確率モデルが，‘偽’であれば，モデルの修正（関数型の修正および誤差項分布型にかんする仮定の変更）なり，推定方法の変更，あるいはデータの変換 etc. を行なってみることが一番の近道となろう。

3. 残差のプロット

仮定３～７をデータに基づいて検討する方法に，Anscombe〔１〕,〔２〕等によって展開された残差分析（residual analysis）がある[6]。これは残差系列を観測不可能な誤差項系列の推定値とみなした上で（この方法についての疑問点を次節で論ずる），誤差項にかんする仮定３～７の違背の検討に役だてようというものである（仮定１の検討にも利用できる）。利用されることの多いダービン＝ワトソン比も，この残差系列から計算される検定統計量であって，一次の系列相関の有無を調べるためのものである[7]。残差をプロットするための方法としては，以下のようなものが考えられる。(1)～(4)の場合，縦軸は残差 e_i である。

(1) 観測時点 i を横軸にとり，点（i, e_i）をプロットする。

(2) 独立変数 x_i を横軸にとり，点（x_i, e_i）をプロットする。

(3) 推定値 $\hat{y}_i$ を横軸にとり，点（$\hat{y}_i$, e_i）をプロットする。

(4) モデルにとり入れられていない新変数 $x_{p+1}=x^*$ を横軸にとり，点（x_i^*,

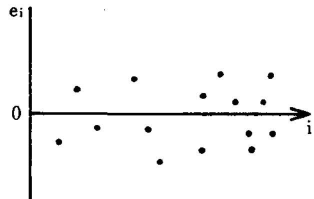
e_i) をプロットする。

(5) e_i を正規確率紙にプッロトする。

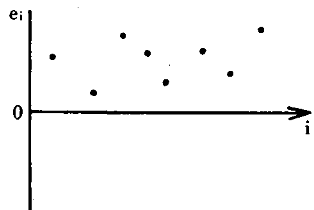
(1)の方法は時系列データを用いる際に最もよく利用され，また非常に有効である。図(3.1)の(a)は正常な状態であり，仮定3，4は満たされていると考えられるが，定数項のないモデルを用いた場合起こり得るのが(b)のようなケースである。定数項のあるモデルを想定している限り，前述の理由により仮定3は常に満たされるが，図は $E(e)>0$ のケースで，モデルに（正の）定数項を含める必要性を教示している。

経済データを用いた場合，頻繁に現われるのが，(c)のような分散増大傾向である。このような，仮定4に対する違背は，加重回帰（weighted regression）[8] 等の操作によって対処することも可能であるが，分散の不均一性（heteroscedasticity）というものは必ずしもデタラメなわけではなく，従属変数や独立変数のサイズに比例的に分散が拡大しているケースの多いことを考慮にいれれば，変数を対数変換することによって分散の不均一を解消できる場合が少なくない。$E(y_i)=e_i$ のとき $V(\log y_i)\fallingdotseq V(y_i)/e_i^2$ となるからである（V は分散を示す）。また，何らかの独立変数 x_k によって，変数を y_i/x_{ki}, $x_{ji}/x_{ki}(j=1, 2, \dots, p)$ と変換することにより分散の不均一性が解消される場合もある。

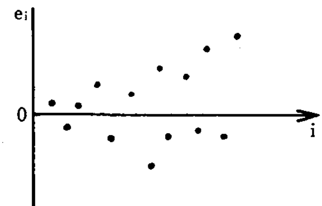
尚，このような散布図を読みとる，と



(a)



(b)



(c)

図 (3.1)

いう方法には客観性に欠けるという欠点があるので，この欠点を補う方法としてよく用いられるものに「連検定（run test)」がある。連検定には様々な形式のものがあるが，ここでは「連の数による検定」を簡単に紹介しておこう。

仮定3により $p_r[e_i>0]=p_r[e_i<0]$ であるから，残差をプラスとマイナス記号で表わせば両者の出現する確率は等しい[9]。連とは同じ記号の続きであり，例えば

$$\underline{+++++}\ \underline{--}\ \underline{+}\ \underline{--}\ \underline{++}\ \underline{---}\ \underline{++}\ \underline{---}$$

では連の数 $r=8$ である。プラス記号の数を m，マイナス記号の数を n で表わすと，連の数が r° 個である確率は次のように表わされる。

$$(3.1)\quad p_r[r=r^\circ]=\begin{cases} \dfrac{2\dbinom{m-1}{k-1}\dbinom{n-1}{k-1}}{\dbinom{m+n}{m}} & r^\circ=2k\ \text{のとき} \\[2ex] \dfrac{\left[\dbinom{m-1}{k-1}\dbinom{n-1}{k-2}+\dbinom{m-1}{k-2}\dbinom{n-1}{k-1}\right]}{\dbinom{m+n}{m}} & r^\circ=2\mathrm{k}-1\ \text{のとき} \end{cases}$$

r の平均と分散は次のようになる。

$$(3.2)\quad \begin{cases} E[r]=\dfrac{2mn}{m+n}+1 \\[2ex] V[r]=\dfrac{2mn(2mn-m-n)}{(m+n)^2(m+n-1} \end{cases}$$

以上の標本分布から，$p_r[r\geqq\bar{r}_\alpha]\leqq\alpha$ および $p_r[r\leqq\underline{r_\alpha}]\leqq\alpha$ なる $\bar{r}_\alpha$（上側確率が α 以下になる最小の点）と $\underline{r_\alpha}$（下側確率が α 以下になる最大の点）が得られるが，実際には既存の数表[10]を用いるか，$m>10$ かつ $n>10$ のときには $\bar{r}_\alpha$，$\underline{r_\alpha}$ が近似的に

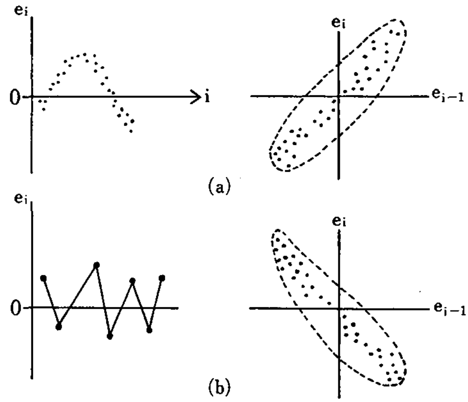
$$(3.3)\quad \bar{r}_\alpha,\ \underline{r_\alpha}=\frac{2mn}{m+n}+1\pm u_\alpha\sqrt{\frac{2mn(2mn-m-n)}{(m+n)^2(m+n-1)}}$$

で与えられることを利用すればよい（u_α は正規分布の片側パーセント点)。

方法(1)は，均一分散性の仮定ばかりでなく，系列相関の有無に関する検討にも用いられる。図（3.2）の(a)は正の系列相関を示すケースで，このような場合，横軸に一期前の残差 e_{i-1} をとり，点（e_{i-1}, e_i）をプロットしてみると右図の

ように右上り傾向となる。これに対して，(b)は負の系列相関のケースで，右の散布図では右下り傾向となる[11]。

系列相関が存在する場合の推定方法としては，「一般化最小二乗法（generalized least squares)[12]」を始め，いくつかのものが考えられているが，ここでは「コクラン＝オーカット



図（3.2）

(Cochran-Orcutt) 法」を紹介するにとどめよう。一次の誤差項系列の相関係数 ρ をとすると，残差系列から ρ を

$$(3.4)\quad \hat{\rho}=\frac{\sum_{i=2}^{n} e_i e_{i-1}}{\sum_{i=2}^{n} e_i^2}$$

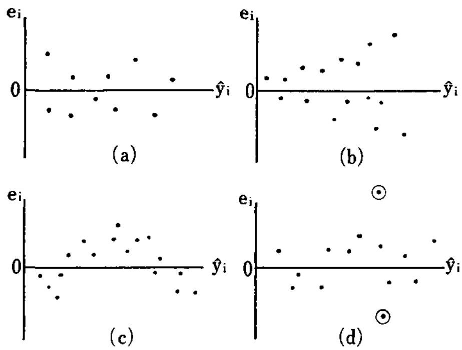
として最小二乗推定し，

$$(3.5)\quad y_i-\hat{\rho}y_{i-1}=\beta_1(x_{1i}-\hat{\rho}x_{1,\,i-1})+\beta_2(x_{2i}-\hat{\rho}x_{2,\,i-1})+\cdots$$
$$+\beta_p(x_{pi}-\hat{\rho}x_{p,i-1})+\varepsilon_i-\hat{\rho}\varepsilon_{i-1},\ i=2,\,3,\,\ldots,\,n$$

に最小二乗法を適用すればよい。

系列相関の発生は，なんらかの有意味な独立変数 x^* がモデルから欠落している場合にも生じやすい。欠落した独立変数が規則的変動を持つ場合，同じ変動パターンが残差系列に現われることが多いものである。このような事態が予測される場合には，プロット法の(4)によって残差系列の規則性の有無について検討し，もしプロットに何らかの規則性が見出されれば，その変数をモデルにとり入れてみるとよい。

プロット法(2)の点（x_i, e_i）のプロットは，独立変数の非線型性を検討する際に用いられるが，独立変数の数が多いときには煩瑣な作業となるので，(3)の方式にしたがって点（$\hat{y}_i$, e_i）をプロットする[13] ことにより，独立変数の非線型



図（3.3）

性，あるいは重要な変数の欠落等を総括的に読みとることができる。図（3.3）において，(a)は正常な場合，(b)は誤差分散の増大傾向が認められる場合，(c)は非線型性の可能性が認められる場合，(d)は異常値（outliers）の存在する場合を示している。

ここで，異常値とは，大半の観測値とくらべて，残差の値が異常に大きい観測値を言う。異常値の検出は，事前に定められた定数 C に対して $|e_i|>C$ となるような残差 e_i に対応する y_i を異常値と判定する。C の定め方には一般論がないが，通常，誤差項の標準偏差の推定値の2～3倍が用いられる。異常値が検出された場合，その原因を調べてみなければならないが，主として質的な構造変化や観測誤差の影響であることが多く，ダミー変数の導入や観測値に適当な修正を施すことによって解決することが可能な場合がある。しかし，経済データの場合，概してその原因が不明であるので，結局異常値を標本から除去することになるが，これは本来，最小二乗法が異常値に対する頑健性が乏しい（異常値によって悪影響を受ける度合が強い[14]）点からして，いたしかたのないところである。

尚，プロット法(3)において，e_i の代りに e_i^2 を使う方法もある[15]が，この方法は変動の大きさを誇張して見せるという長所の反面，マイナスの値をとる残差もその平方として捉えるため，例えば分散の増大傾向と減少傾向の区別が不可能になるという欠点を持つ。

方法(5)は誤差項分布の正規性にかんして検討するためのもの[16]で，正規確率紙上に残差が直線的に並ぶかどうかで判定を行なうものである。推定値の計算のみに関心がある場合には，正規性の仮定は必要ないので行なわない。

最後に，以上のような残差のプロットによって仮定の検討を行なう場合，統計的な判断のみならず，経済理論および現実の経済にかんする十分な知識を持

ち合わせていることが不可欠であることを付け加えておこう。

4. 残差の標準化と Theil の BLUS 残差

前節の冒頭で触れておいた疑問点について考えてみよう。モデル (2.1) において，仮定により $E(\boldsymbol{\varepsilon})=\mathbf{0}$, $Cov(\boldsymbol{\varepsilon})=\sigma^2\boldsymbol{I}$ であるが，これをもとに残差の期待値と分散を求める。$\boldsymbol{y}=\boldsymbol{X}\hat{\boldsymbol{\beta}}+\boldsymbol{e}$ であるから，残差平方和は

$$(4.1)\quad \|\boldsymbol{e}\|^2=\boldsymbol{e}'\boldsymbol{e}=(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})$$
$$=\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{y}$$

となり，$\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ は巾等行列であるから，

$$(4.2)\quad \boldsymbol{e}=[\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\boldsymbol{y}=[\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\boldsymbol{\varepsilon}$$

である。$\boldsymbol{M}=\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ とおいて期待値をとれば，

$$(4.3)\quad E(\boldsymbol{e})=\boldsymbol{M}E(\boldsymbol{\varepsilon})=\mathbf{0}。$$

残差分散は

$$(4.4)\quad Cov(\boldsymbol{e})=E(\boldsymbol{e}\boldsymbol{e}')=\boldsymbol{M}E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\boldsymbol{M}=\boldsymbol{M}\sigma^2\boldsymbol{I}\boldsymbol{M}$$
$$=\sigma^2\boldsymbol{M}$$

となる。(4.2) および $\boldsymbol{e}$ の分散共分散行列が特異である[17] ことから，仮定7の $\boldsymbol{\varepsilon}\sim N(\boldsymbol{O},\sigma^2\boldsymbol{I})$ のとき $\boldsymbol{e}$ は特異な分散共分散行列 $\sigma^2\boldsymbol{M}$ をもった多変量正規分布に従う。したがって，標準モデルの仮定のもとでも e_i の分散は互いに異なり，e_i と e_j の相関もゼロとならない。かくして e_i を ε_i の推定値とした前節の残差プロット法には疑問が生じたことになる。

$(\boldsymbol{X}'\boldsymbol{X})^{-1}$ の第 (i, j) 要素を $(\boldsymbol{X}'\boldsymbol{X})^{-1}_{ij}$, $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ の第 (i, j) 要素を W_{ij} と書くことにすれば

$$(4.5)\quad W_{ij}=\sum_{k=1}^{p}\sum_{h=1}^{p}x_{ik}(\boldsymbol{X}'\boldsymbol{X})^{-1}_{ij}x_{jh}$$

となり，$p/n\to 0$ のとき $W_{ij}\to 0$ である[18]。それ故，$\sigma^2\boldsymbol{M}=\sigma^2[\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']$ における行列 $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ の各要素は比較的小さい値であり，特に $p/n\gg 0$ のときには $Cov(\boldsymbol{e})=\sigma^2\boldsymbol{I}$ となる。その意味では **p/n が十分に大きいときには前節で述べたような残差の系列をそのまま使って差しつかえない**。しかし，より正確には

(4.6) $\dfrac{e_i}{s\sqrt{1-W_{ii}}}$

を用いれば結果は改善されるが，概して $\sqrt{1-W_{ii}}$ は1に近い値であることから，次式のように残差を

(4.7) $e'=\dfrac{e_i}{s}$

標準化して用いることが多い。s は母分散の不偏推定値 $s^2=\dfrac{ee'}{n-p}$ の平方根である。

Theil〔12〕,〔13〕は上述のような最小二乗残差の分散が定数行列とならないことから生ずる問題点を解決するべく, "BLUS (best linear unbiased scalar-covariance) 残差" なるものを提唱している。これは前節の仮定1～7に加えて

仮定8　残差系列が無相関で，かつ分散一定 ($Cov(\boldsymbol{e}^*)=\sigma^2\boldsymbol{I}$)

を付加した上で，最小二乗残差 $\boldsymbol{e}^*$ を求めようというものである。

$\boldsymbol{A}'$ を $(n\times p)\times n$ の任意の定数行列とし，$\boldsymbol{e}^*=\boldsymbol{A}'\boldsymbol{y}$ と記す。$Cov(\boldsymbol{e}^*)=E(\boldsymbol{A}'\boldsymbol{\varepsilon\varepsilon}'\boldsymbol{A})=\sigma^2\boldsymbol{A}'\boldsymbol{A}$ が $\sigma^2\boldsymbol{I}$ に一致するためには

仮定9　$\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{I}$

が必要となる。この仮定のもとでは $\boldsymbol{e}^*$ の次数は $(n-p)\times n$ となるため，最終的には $n-p$ 個の残差しか得られない(19)。p 個の系列を $n\times n$ 行列から除くことによってできる $(n-p)\times n$ 次の単位行列を $\boldsymbol{J}'$ で表わすと，$\boldsymbol{e}$ のうちで推定されるべき小ベクトルは $\boldsymbol{J}'\boldsymbol{e}$ で示され，

(4.8) $\boldsymbol{J}'=[\underset{(n-p)\times p}{\boldsymbol{O}}\ \vdots\ \underset{(n-p)\times(n-p)}{\boldsymbol{I}}]$

に対応して

(4.9) $\boldsymbol{X}'=[\underset{p\times p}{\boldsymbol{X}_0'}\ \vdots\ \underset{p\times(n-p)}{\boldsymbol{X}_1'}]$

と分割すると，仮定9によって $\boldsymbol{X}_0'$ が非特異であることが要求される。

最小二乗基準 $E[\boldsymbol{e}'(\boldsymbol{A}-\boldsymbol{J})(\boldsymbol{A}-\boldsymbol{J})'\boldsymbol{e}]\to\min$ を満たす $\boldsymbol{A}'$ を条件 $\boldsymbol{A}'\boldsymbol{x}=\boldsymbol{0}$(20) および仮定9のもとで求めると，$\boldsymbol{A}'$ は次の2式

(4.10) $(\boldsymbol{M}-\boldsymbol{A}\boldsymbol{A}')\boldsymbol{J}=\boldsymbol{0}$

(4.11)　$A'J=J'A$

を同時に満たす行列でなくてはならない。

(4.10) および (4.11) を同時に満足するような A' を具体的に計算するためには，e のなかで推定されない p 個の項が初めから順に p 個並んでいなければならないが，これは観測値を並び変えることによって，一般性を失うことなく常に可能である。(4.8) および (4.9) に対応して，M と A を

$$(4.12)\quad M=\begin{array}{c}\begin{array}{cc} \scriptstyle p\times p & \scriptstyle p\times(n-p) \end{array}\\ \left(\begin{array}{c:c} M_{00} & M_{01} \\ \hdashline M_{10} & M_{11} \end{array}\right)\\ \begin{array}{cc} \scriptstyle (n-p)\times p & \scriptstyle (n-p)\times(n-p) \end{array}\end{array},\quad A=\begin{array}{c}\scriptstyle p\times p\\ \left(\begin{array}{c} A_0 \\ \hdashline A_1 \end{array}\right)\\ \scriptstyle (n-p)\times(n-p)\end{array}$$

と分割する。添字 $_0$ は最初の p 個の行または列，添字 $_1$ は残りの $n-p$ 個の行または例を示している（A_1 は非特異行列であることに注意)。ここで

$$(4.13)\quad \begin{cases} M_{00}=I_p-X_0(X'X)^{-1}X_0' \\ M_{01}=-X_0(X'X)^{-1}X_1' \\ M_{10}=M'_{01} \\ M_{11}=I_{n-p}-X_1(X'X)^{-1}X_1' \end{cases}$$

なる関係式から，次の3式が (4.10) および (4.11) と同値となる。

$$(4.14)\quad \begin{cases} M_{01}=A_0A_1' \\ M_{11}=A_1A_1' \\ A_1=A_1' \end{cases}$$

M および M の部分行列は観測値から直接計算可能であるから A_1' を得，その結果を $e^*=A'y=\begin{bmatrix}A_0\\A_1\end{bmatrix}'y$ に代入して，$e^*_i(i=p+1,\ p+2,\dots,\ n)$ を求めることができる。

以上のように BLUS 残差は計算上厄介な統計量であるが，残差の分析には前述の e や e' を用いるよりも望ましいことは言うまでもない。BLUS 残差系列の導出にはコンピュータの利用[21] が不可欠であるが，コンピュータを用いても相当に煩瑣な仕事であることを覚悟して取り組まねばならない。残差プロットの方法は手軽さがその長所のひとつであるから，一般には e' を（あるいは p/n が十分大きければ e を）使用する方法で十分であろう。

注(1)　「計量モデル (econometric model)」と呼んでもよい。

(2)　佐和〔9〕86ページ。

(3)　他にもいくつかの仮定が用いられるが，ここでは触れないことにする。例えば x_i は確定変数である，等々。

(4)　最小二乗推定量は (2.1) における残差ベクトルの norm, $\sum_{i=1}^{n}\varepsilon_i=\|\boldsymbol{\varepsilon}\|^2$ を最小化するパラメータ $\boldsymbol{\beta}$ の値として定義される。$\partial\|\boldsymbol{\varepsilon}\|^2/\partial\boldsymbol{\beta}=\mathbf{0}$ を解いて $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}=\boldsymbol{X}'\boldsymbol{y}$ を得るが，これを正規方程式 (normal equation) と呼び，その解 $\hat{\boldsymbol{\beta}}=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ が最小二乗推定量である。

(5)　実際問題として，rank$(\boldsymbol{X})<p$, すなわち $\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_px_p=0$ となるようなすべてがゼロでない定数 $\alpha_1, \alpha_2, \ldots, \alpha_p$ が存在するようなことはまずなく，現実におこり得るのは $\alpha_1x_1+\alpha_2+x_2+\cdots+\alpha_px_p\fallingdotseq0$ のようなケースであり，パラメータの推定値そのものは計算できても推定量分散は非常に大きくなる。尚，前者の等号の場合を Johnston〔6〕は「完全な多重共線関係」と呼んでいる。Johnston〔6〕邦訳193ページ。

(6)　Anscombe = Tuckey〔2〕141-160p. 残差分析は Draper = Smith〔4〕邦訳88—106ページ，奥野他〔8〕92—112ページに詳しい。

(7)　ダービン＝ワトソン比は $d=\dfrac{\sum_{i=2}^{n}(e_i-e_{i-1})^2}{\sum_{i=1}^{n}e_i^2}$ で定義される。奥野他〔8〕109ページ，Gujarati〔5〕235p.

(8)　佐和〔9〕134ページ，佐和〔10〕124ページ，Gujarati〔5〕198p. を参照のこと。

(9)　$e_i=0$ なる残差の出現は実際問題として考慮する必要はない。

(10)　統計数値表編集委員会員編「統計数値表コンサイス版」日本規格協会，1977，82—3ページ。又は Draper = Smith〔4〕邦訳 100—101ページ。

(11)　系列相関の有無を残差のプロットで検討する方法は Gujarati〔5〕219-235p. に詳しい。

(12)　Johnston〔6〕邦訳 173〜179ページ。佐和〔9〕99ページ。

(13)　残差 e_i と観測値 y_i との間には相関がある $(r^2(\boldsymbol{y}, \boldsymbol{e})=1-R^2)$ ので，点 (y_i, e_i) のプロットはこのような検討には役立たない。

(14)　佐和〔10〕133—134ページを参照されたい。

(15)　Gujarati〔5〕201-203p.

(16)　Daniel = Wood〔3〕27-32p. に詳しい。

(17)　$tr\boldsymbol{M}=tr\boldsymbol{I}_n-tr\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'=n-tr(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}=n-tr\boldsymbol{I}_p=n-p$ であるから，rank$\boldsymbol{M}=n-p$ となり，$Cov(\boldsymbol{e})=\sigma^2\boldsymbol{M}$ も特異行列となる。

(18)　佐和〔10〕131ページを参照されたい。

(19)　どの p 個の残差を除外するかを一般的に論ずることは困難であるが，Theil〔12〕は数値例をとりあげ，その中で最初と最後の合計 p 個の残差を除外している。Theil

〔12〕1070p.

(20) $e^* = A'y$ より $A'y = A'X\beta + A'e$ であるから，推定における誤差は $A'X\beta + A'e - e$ となり，この式がゼロベクトルとなるための必要十分条件は $A'X=0$ である。Theil〔12〕1068p.

(21) コンピュータによる残差プロットでは標準化された残差 e_i/s を用いることが多い。

文献

〔1〕 Anscombe, F. J., *The rejection of outliers,* 1960 Technometrcs 2.

〔2〕 Anscombe, F. J., and J. W. Tuckey, *The examination and analysis of residuals,* 1963, Technometrics 5.

〔3〕 Daniel, C. and F. S. Wood, *Fitting Equations To Data : Computer Analysis of Multifactor Data* 2nd ed. 1980, New York.

〔4〕 Draper, N. R. Smith, *Applied Regression Analysis,* 1966（中村慶一訳「応用回帰分析」1967，森北出版）

〔5〕 Gujarati, D., *Basic Econometris,* 1978, New York.

〔6〕 Johnston, J., *Econometric Methods,* 1963（竹内啓訳「計量経済学の方法」東洋経済新報社，1964）

〔7〕 Kennedy, P., *A Guide To Econometrics,* 1979, New York.

〔8〕 奥野＝久米＝芳賀＝吉澤「多変量解析法」1971，日科技連

〔9〕 佐和隆光「計量経済学の基礎」1970，東洋経済新報社

〔10〕 佐和隆光「回帰分析」1979，朝倉書店

〔11〕 Theil, H. and A. L. Nagar, *Testing the Independence of Regression Disturbances,* Journal of American Statistical Association (1965) 60.

〔12〕 Theil, H., *A Simplification of the BLUS Procedure for Analyzing Regression Disturbances,* Journal of American Statistical Association (1968) 63.

〔13〕 Theil, H., *Principles of Econometrics,* 1971, New York.