第４章　線形モデル：
生産関数と消費関数

1　はじめに

本章では経済学の実証研究でもっとも頻繁に使われる線形モデルについて解説する。線形モデルとは一般に被説明変数 y を複数の説明変数 $\mathbf{x}$ によって説明する $y = \alpha + \mathbf{x}'\beta + \varepsilon$ のようなモデルを指している。このモデルの意味していることは、(1) 説明変数 $\mathbf{x}$ によって y を条件付予測している、あるいは (2) 説明変数 $\mathbf{x}$ は y の原因である、ということであるが、(1) と (2) は厳密には違う。すなわち、(1) では相関関係があることを利用して、y を推測できるということであるのに対して、(2) では $\mathbf{x}$ が y の原因であるという因果関係にまで踏み込んだ解釈をしている。言うまでもないが、(1) の方が (2) よりも弱い関係を示唆している。逆に言えば、因果関係があるというためには、y に影響を与えそうな全ての変数が $\mathbf{x}$ に含まれており、その他に y に影響を与えるのは純粋な誤差 ε だけであるという条件および説明変数 $\mathbf{x}$ 間の相関や副次的な因果関係が排除されているという条件が必要になってくる。これらの厳密な因果関係の分析はプログラム評価あるいは政策評価という分野で行われており、本書でも第 12 章で触れるつもりである。

本章ではより緩やかな線形関係、すなわち、説明変数 $\mathbf{x}$ の変動が y の変動を統計的に有意な水準で説明できるかどうかを考えていく。

線形モデルに関してもう一つ注意しておくべき点は純粋に線形関係にあると考えられるモデルと、本質的には非線形関係なのだが、局所的に線形近似して扱っているモデルは別だということである。実際、第 3 部で扱うように、ミクロ計量経済学の問題の多くは非線形モデルとして定式化されており、線形モデルとして扱われていても非線形モデルを近似したものであることが多い。これも第 3 部で明らかになるが、線形モデルであれば、推計パラメータ β は $\mathbf{x}$ の値に関わらず一定であるが、非線形モデルであれば、その曲線の傾きを示すパラメータ β は、$\mathbf{x}$ の位置に応じて変化し、政策的に一般的な評価を与えることが極めて難しいことも知られている[1]。そのために政策に用いるマクロモデルなどでは本質的には非線形モデルを対数線形化などして用いて

[1] 非線形モデルでは推定されたパラメータ β を直接比較することは、あまり意味がないので、限界効果を見たりオッズ比を計算して比較する。詳しくは第 3 部を参照されたい。

いる。

さらに、ミクロデータを使う場合と金融マクロ時系列データを使う場合には、おなじ線形モデルを推定するにしても、その扱いが違ってくることにも注意すべきである。金融マクロ時系列データであれば、同一の金融変数であったり、集計されたマクロ変数であったりするので、データの中に異質性が含まれていたり、誤差項の分散が不均一になるような状況は、マクロショックとして扱うことが可能である。クロスセクションデータやパネルデータの形態をとるミクロデータでは異質性の要因が、個別主体のもともとの属性・趣向の違い、個別経済主体への独立したショック、地域や特定グループへのショック、あるいはマクロ経済全体へのショックなど様々であり、それらの要因を注意深く選り分けながら、共通したパラメータを推定する必要が出てくる。

2 線形モデルの特性

次のようなクロスセクション・データがあるとしよう。サンプル数 N、被説明変数 $y(N \times 1$ 行列)、説明変数のベクトル $\mathbf{x}(N \times K$ 行列) とすれば、線形モデルは一般に次の様に表現できる。

$$y = E(y|\mathbf{x}) + \mathbf{u}$$

ここで $E(y|\mathbf{x})$ は y の条件付期待値を表しており、$\mathbf{u}(N \times 1$ 行列) は誤差項である。

この条件付期待値の部分を次のように特定化して線形回帰モデルが定義される。

$$y = \mathbf{x}'\beta + \mathbf{u}$$

パラメータ β の推定方法としては誤差項を何らかの意味で最小にするようにするという考え方を使う。

最も一般的なのは誤差二乗を最小にするというもので、**最小二乗法**（Ordinary Least Squares:OLS）として知られている ($\min \mathbf{u}^2$)。ここでは平均値からの誤差を測っている。また誤差項の絶対値を最小にするという考え方、**絶対誤差最小法**（Absolute Error Loss Minimization：$\min |\mathbf{u}|$）と呼び、一般には中位値（median）からの距離を測る。これら 2 つの方法は誤差が正であれ負であれ区別していないが、過少推定と過大推定に違うペナルティを課す**非対称絶対誤差最小法**（Asymmetric Absolute Loss Minimization:$\min(1-\alpha)|\mathbf{u}|\ if\ \mathbf{u} < 0,\ \min \alpha|\mathbf{u}|\ if\ \mathbf{u} \geqq 0$）も考えられる。このパラメータ α が 0.5 であれば、絶対誤差最小法に収斂するが、それ以外の値をとれば非対称になる。これは区分平均からの誤差を用いる**区分回帰法**（Quantile Regression）として知られている。

2.1 最小二乗法

ここでは定数項を入れた線形回帰モデルを次のように定義しよう。

$$y = \alpha + \mathbf{x}'\beta + \mathbf{u}$$

最小二乗法の問題は次のようになる。

$$\min \mathbf{u}^2 = \min(y - \alpha - \mathbf{x}'\beta)$$

この解の一階条件は $-2E(\mathbf{u}) = 0$ と $-2E(\mathbf{xu}) = 0$ となり、これは $E(\mathbf{u}) = 0$ と $Cov(\mathbf{x}, \mathbf{u}) = 0$ を意味している。この結果を用いてパラメータは次のように推定される。

$$\beta = (Var(\mathbf{x}))^{-1}Cov(\mathbf{x}, y) = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'y$$
$$\alpha = E(y) - E(\mathbf{x})'\beta$$

ここで $Var(\mathbf{x})$ は $\mathbf{x}$ の分散であり、$\mathbf{x}'\mathbf{x}$ の逆数が存在するとき、それを**非特異（**nonsingular) と呼ぶ　。$Cov(\mathbf{x}, y)$ は $\mathbf{x}$ と y の共分散である。最小二乗法が**最良線形不偏推定量**（Best Linear Unbiased Estimator: BLUE）となるためには、$E(\mathbf{u}|\mathbf{x}) = 0$ が成り立つことが必要である。これが成り立つことによって y の期待値は $E(y|\mathbf{x}) = \alpha + \mathbf{x}'\beta$ となり、確率極限の推定量 β は不偏 (unbiased) になる[2]。

線形モデルがデータに対してどれぐらい説明力があるかどうかの Goodness of Fit は**決定係数**（Coefficient of Determination）、一般には R^2 と呼ばれている統計量を用いることが多い。考え方は簡単で、平均からのバラつきの二乗和を説明変数で説明できる部分の二乗和と出来ない部分の二乗和にわけて、説明変数で説明できる部分のシェアを求めたものである。すなわち次のように定義できる。

$$R^2 = 1 - \frac{\sum_{t=1}^{n} \widehat{u}_t^2}{\sum_{t=1}^{n}(y_t - \overline{y})^2}$$

また、自由度を修正した決定係数は次のように定義できる。

$$adjR^2 = 1 - \frac{\frac{1}{n-k}\sum_{t=1}^{n} \widehat{u}_t^2}{\frac{1}{n-1}\sum_{t=1}^{n}(y_t - \overline{y})^2}$$

ここで k は説明変数の数をあらわす。一般には**自由度修正済み決定係数**を用いる方が望ましい。

[2] ここでは最小二乗法の理論についてこれ以上詳しい説明を行わないが、興味のある方は Amemiya (1985) や畠中（1991）の該当箇所を参照されたい。

もう一つ線型モデルの個別パラメータが統計的に有意かどうかは t 検定を行うのが標準的である。**t 統計量**は次のように定義できる。

$$t_{\beta_i} = \frac{\widehat{\beta}_i - \beta_{i0}}{(V(\widehat{\beta}_i))^{1/2}}$$

ここで $\widehat{\beta}_i$ は推定パラメータ、β_{i0} は理論的に想定されるパラメータ値（多くの場合 0)、$V(\widehat{\beta}_i)$ は推定パラメータの分散である。この $(\widehat{\beta}_i - \beta_{i0})$ が漸近的に正規分布に従い、$Var(\widehat{\beta}_i)$ が漸近的に自由度 $n-k$ のカイ二乗分布に従えば、その比である t_{β_i} は漸近的に**スチューデントの t 分布**に従うことが明らかにされている。それぞれのパラメータに対して t 統計量を計算し t 分布に基づいて検定することが必要である。

2.2 不均一分散

ミクロ計量経済学では誤差項は**条件付き不均一分散**（conditionally heteroskedastic）となることが多い。これは $V(u_i|x_i) = E(u_i^2|x_i) = \sigma_i^2$ ということで誤差項の分散が個人毎に異なり、それが説明変数である条件 x_i に依存しているという意味である。

では、この不均一分散は、どのように発見でき、検定できるだろうか。Breusch and Pagan (1979) は次のようなラグランジェアン乗数法検定を提案した。

$$\widehat{u}_i^2 = d_1 + d_2 z_{i2} + d_3 z_{i3} + d_4 z_{i4} + ... + d_l z_{il} + v_i$$

帰無仮説は $\widehat{u}_i^2$ は一定である、ということで、具体的には $d_j = 0 \quad j = 2, 3, 4, ...l$ を検定する。ここで用いた説明変数 $\mathbf{z}_j$ はもともとのモデルから推定された $\widehat{y}_i$ でもいいし、それ以外の変数でもかまわない[3]。

基本的な考え方としては、不均一分散が強く存在することが確認されれば、それはモデルの定式化に問題があるか、データに複数の違った行動様式に従っている主体が含まれているかであり、何らかの対応をすることが考えられる。正当な対策としては、見落としている要因がないかどうかモデルを再検討することである。

ただ、モデルを変更するほどではないにしても、不均一分散の存在が、推定係数の有意性検定にバイアスをもたらす可能性がある。White (1980) は分散共分散行列 $\mathbf{M_{x\Omega x}} = p\lim N^{-1}\sum_{i=1}^{N} u_i^2 x_i x_i'$ の推定においてバイアスを取り除く方法として、u_i を最小二乗法推定誤差 $\widehat{u}_i = y_i - x_i'\widehat{\beta}$ を用いて分散共

[3]STATA ではこの検定は Breusch-Pagan/Cook-Weisberg test for heteroskedasticity と呼ばれ、hettest というコマンドで検定量が計算される。変数 $\mathbf{z}$ の選択に関してはルールがあるわけではない。Davidson and MacKinnon (2004, p.269) 参照。

分散行列を計算することを提案した。漸近的に $\hat{u}_i \to u_i$ となれば次の分散共分散行列は不偏推定となる。

$$\widehat{\mathbf{M}}_{\mathbf{x\Omega x}} = N^{-1}\textstyle\sum_{i=1}^{N}\hat{u}_i^2 x_i x_i' = N^{-1}\mathbf{x}'\hat{\Omega}\mathbf{x}$$

ここで $\hat{\Omega} = Diag(\hat{u}_i^2)$ である。これに $\widehat{\mathbf{M}}_{\mathbf{xx}} = N^{-1}\mathbf{x}'\mathbf{x}$ を仮定して、最小二乗法推定パラメータ β の分散は次のように表せる。

$$\begin{aligned}\hat{V}(\hat{\beta}_{OLS}) &= (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\Omega\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\\ &= (\textstyle\sum_{i=1}^{N}x_i x_i')^{-1}\textstyle\sum_{i=1}^{N}\hat{u}_i^2 x_i x_i'(\textstyle\sum_{i=1}^{N}x_i x_i')^{-1}\end{aligned}$$

これが White(1980) が提案した推定パラメータの分散の推定の仕方であり、これを用いて計算した標準誤差を**不均一分散頑強標準誤差**（heteroskedasticity-robust standard error）と呼ぶ。ミクロデータを用いる場合、ほとんどのケースで不均一分散の問題に直面するが、この不均一分散頑強標準誤差を用いれば、推定パラメータの標準誤差は不偏に推定できる。従って、ミクロ計量経済学ではほぼ常にこの頑強標準誤差を計算し、**頑強 t 統計量**を計算することが望ましい。

2.3 加重最小二乗法

不均一分散に対応した標準誤差の推定だけではなく、モデルの推定方法で不均一分散によるバイアスを取り除く方法も考えられている。

次のような最も簡単な線形モデルを考えよう。

$$\begin{aligned}y &= \mathbf{x}'\beta + \mathbf{u}\\ E(\mathbf{uu}') &= \Omega\end{aligned}$$

誤差項の共分散行列を Ω とする。Ω は非特異（nonsingular）である。もし $\Omega = \sigma^2\mathbf{I}$ であれば、誤差項は独立で均一分散であり、最小二乗法が最良線形不偏推定量をもたらす。しかし、$\Omega \neq \sigma^2\mathbf{I}$ であれば、誤差項は不均一分散であり、共分散項もゼロではない。この場合、もともとの線形モデルに次のように定義される行列（Ψ'）を掛けてモデルを変換してみる。

$$\Omega^{-1} = \Psi\Psi'$$

すなわち、

$$\Psi' y = \Psi'\mathbf{x}'\beta + \Psi'\mathbf{u}$$

この式を最小二乗法推定すれば、パラメータは次のように表せる。

$$\beta_{GLS} = (\mathbf{x}'\Psi\Psi'\mathbf{x})^{-1}\mathbf{x}'\Psi\Psi' y = (\mathbf{x}'\Omega^{-1}\mathbf{x})^{-1}\mathbf{x}'\Omega^{-1} y$$

この推定量を**一般化最小二乗法**（Generalized Least Squares: **GLS**）推定とよぶ。変換された誤差行列は単位行列（**I**）になる。

$$\begin{aligned} E(\Psi'\mathbf{u}\mathbf{u}'\Psi) &= \Psi' E(\mathbf{u}\mathbf{u}')\Psi = \Psi'\Omega\Psi \\ &= \Psi'(\Psi\Psi')^{-1}\Psi = \Psi'(\Psi')^{-1}\Psi^{-1}\Psi = \mathbf{I} \end{aligned}$$

誤差項の共分散行列は実際に線型モデルを最小二乗法推定して得られたものを用いるしかない。すなわち $\widehat{\Omega} = \Omega(\widehat{\beta})$ を用いてパラメータ β を推定する方法を**実行可能一般化最小二乗法**（Feasible Generalizaed Least Squares: **FGLS**）と呼ぶ。

$$\beta_{FGLS} = (\mathbf{x}'\widehat{\Omega}^{-1}\mathbf{x})^{-1}\mathbf{x}'\widehat{\Omega}^{-1} y$$

これらの推定方法は明らかに最小二乗法より効率的（最小分散）である。

これまでの議論では誤差項の共分散行列 Ω は既知であるか、あるいは、漸近的に真の値に近づくことを仮定してきた。しかし、それが一般には保証されない場合には、ある変数の分散（例えば Σ とする）をウェイトとして推定する**加重最小二乗法**（Weighted Least Squares:**WLS**) を考えることも有益である。

$$\widehat{\beta}_{WLS} = (\mathbf{x}'\widehat{\Sigma}^{-1}\mathbf{x})^{-1}\mathbf{x}'\widehat{\Sigma}^{-1} y$$

ここでは一般には $\Sigma \neq \Omega$ である。言うまでもなく、WLS は GLS や FGLS を特殊ケースとして含む一般表現であると考えることができる。

経験則として不均一分散の問題がある時には、最初に想定したモデルを規模の大きな変数、例えば人口や GDP などで割って人口比や GDP 比で表すことによって、誤差項の不均一分散の程度が大幅に改善されることが知られている。これは、加重最小二乗法の簡便版である。

2.4 区分回帰法

サンプルの分布が正規分布のように対称分布ではなく、非対称であれば、平均値と中位値は異なってくる。最小二乗法では平均値からの乖離である誤差の二乗を最小にすることでパラメータを推定する方法であるが、サンプルの大半が平均値周辺に分布していれば、この方法によって推定されたパラメータはサンプル集計値として扱ってもいいが、大半が平均値から離れた中位値

あるいはその他の区分に分布している場合には、中位値あるいはその他の区分平均値からの乖離を誤差として計測し、それに基づいてパラメータを推定することも意味があるだろう。このような推定方法を**区分回帰法**（Quantile Regression）と呼ぶ[4]。

連続確率変数 y の第 q 区分の閾値 μ_q を定義し、y が μ_q より小さい値をとる確率を q とする。

$$q = \Pr(y \leq \mu_q) = F_y(\mu_q)$$

ここで F_y は y の累積密度関数である。逆関数は次のように表せる。

$$\mu_q = F_y^{-1}(q)$$

ここで $y = \mathbf{x}'\beta + u$ とすると、上式は次のように書き換えることができる。

$$\mu_q(\mathbf{x}) = F_{y|\mathbf{x}}^{-1}(q)$$

区分回帰法で第 q 区分のパラメータ β_q を求めるということは、次式を β に関して最小化することである。

$$Q_N(\beta_q) = \sum\nolimits_{i:y_i \geq \mathbf{x}_i'\beta}^{N} q|y_i - \mathbf{x}_i'\beta_q| + \sum\nolimits_{i:y_i < \mathbf{x}_i'\beta}^{N} (1-q)|y_i - \mathbf{x}_i'\beta_q|$$

これは先に論じた**非対称絶対誤差最小法**の一種である[5]。

3 生産関数と消費関数の推定

ミクロ経済学およびマクロ経済学の教科書の両方に出てくる典型的な経済モデルとして生産関数と消費関数がある。本節ではこの二つの経済モデルを推計してみよう。ここでは前節までに学んだ推定方法を実際に使ってみることを目的としており、最先端の生産関数や消費関数を推計することは目的としていない。また、ここではマクロ時系列データによる消費関数からミクロ・クロスセクションデータによる消費関数のどこが違うのかをデータから直感的に理解していただきたい。

ここで用いたデータは生産関数に関しては Winklemann and Boes (2005) で用いられ、ホームページで公開されているものを使った（http://www.sts.uzh.ch/data/cobb.html）。データは人為的に作られたとされている。また消費関数に関してはマクロ時系列データ (1959-95) は Wooldridge (2003) で用いられているもので、彼のホーム

[4]区分回帰法の基本文献は Koenker (2005) である。この手法の詳細については Koenker(2005) を参照されたい。

[5]この最適化問題は微分可能連続関数ではないので、線形計画法を用いて解く。

ページで公開されているものを用いた (http://www.msu.edu/˜ec/faculty/wooldridge/book2.htm)。ミクロクロスセクションデータは世界銀行が Living Standard Measurement Study の一環として 1997 年にベトナムで行った調査（5006 サンプル）を用いている。このデータは Cameron and Trivedi (2005) が使い、彼らのホームページで公開しているものである (http://cameron.econ.ucdavis.edu/mmabook/mma.html)。

3.1 生産関数

生産関数の中で最も基本的なものはコブ・ダグラス型生産関数と呼ばれているものである。具体的には次のようなモデルである。

$$y = F(K, L) = AK^{\alpha}L^{\beta}$$

ここで y は生産量（あるいは生産総額）、A は技術水準を表す定数、K は資本ストック量（額）、L は労働投入量（労働者数あるいは労働時間 $\times$ 賃金）である。ここでパラメータ $\alpha + \beta$ が 1 より小さいか、大きいか、等しいかで生産活動の規模の経済性を見ることができる。すなわち、$\alpha + \beta > 1$ であれば収穫逓増、$\alpha + \beta = 1$ であれば収穫一定、$\alpha + \beta < 1$ であれば収穫逓減という。

実際に推定するモデルは上式の対数をとり、統計的な誤差を想定して次のように定式化できる。

$$\ln y = \ln A + \alpha \ln K + \beta \ln L + u$$

推定結果は表 1 に掲載されている。本章で論じた OLS、WLS、GLS に加えて、最尤法による一般線形モデル（Genaralized Linear Model: GLM)）でも推定した。OLS 推定では不均一分散検定で若干の不均一性が見出されるので、WLS、GLS では lnk 資本ストックの分散をウェイトとして使い、推定した。しかし、全体としてはパラメータ値も頑強標準誤差に基づく t 値もそれほどかわらなかった。表 1 にある Ramsey RESET test とは脱落変数（omitted variable）の検定であり、モデル特定化に誤りがないかどうかを簡便に検定できるものである[6]。ここではモデル特定化に問題はないという帰無仮説が棄却されていない。

さらに規模の経済性に関して収穫一定 $\alpha + \beta = 1$ であるかどうかを F 検定した。その結果、全ての推定で棄却され、収穫逓減型の生産関数 $\alpha + \beta < 1$ であることがわかった。図 1 で $(x, y) = (K/L, y/L)$ を描くと、労働者 1 人当たりの資本ストックが増加するにつれて労働者 1 人当たりの生産量は逓減していることが直感できるだろう。

[6] 具体的には、$y = \mathbf{x}'\beta + u$ がもともとのモデルだとして、これに変数 $\mathbf{z}$ を加えて、$y = \mathbf{x}'\beta + \mathbf{z}'t + v$ を推定し、パラメータ $t = 0$ かどうかを検定する。ここで $\mathbf{z}$ としては、被説明変数 y の推定値の 2 乗項、説明変数 $\mathbf{x}$ の 2 乗項などが用いられる。

3.2 消費関数

消費関数の基本的なモデルは可処分所得で消費を説明するケインズ型消費関数か異時点間での消費の平準化をモデル化したライフサイクル恒常所得仮説に基づく消費関数である。本節では以下のような 3 つのモデルを推定した。

$$c_i = a + by_i + u_i$$
$$\ln c_i = d + e \ln y_i + fi3 + v_i$$
$$\triangle \ln c_i = g + h \triangle \ln y_i + \varepsilon_i$$

ここで c は一人当たりの実質消費、y は一人当たり実質可処分所得、$i3$ は 3ヶ月物財務省短期証券利回り、$\triangle$ は一階の階差を指し、$\triangle \ln c = gc$ は一人当たり実質消費の成長率、$\triangle \ln y = gy$ は一人当たり実質可処分所得の成長率を意味している。推定結果は表 2 に載せてある。

第 1 式は単純なケインズ型消費関数で決定係数も 0.997 と高く、可処分所得にかかる係数も 0.779 で有意である。不均一分散の問題もなさそうだが、モデル特定化としては他に追加すべき変数を落としている可能性が棄却できない。第 2 式は対数をとった変数と金利データを加えた。このモデルも決定係数が 0.998 と高く、可処分所得にかかる係数は 0.956 と極めて高く有意に出ている。また金利の係数は-0.002 と負で有意であり、理論的にも整合的な結果である。これら 2 本の推計式で用いたデータは図 2 に描かれている。ともに一人当たり実質消費のレベルであれ、その対数表示であれ、一人当たり実質可処分所得のレベルおよびその対数表示と見事な線形関係をもっていることがわかる。

第 3 式ではライフサイクル恒常仮説消費関数が正しければ、可処分所得成長率の係数 h は 0 になるはずである。すなわち、消費平準化を行っている限り、前期と今期の実質消費の階差によって説明できない差は予想外のショック以外にはないと考えられるからである。実際に表 2 を見ると、一人当たり実質可処分所得成長率の係数は 0.571 で有意に効いている。このことは、少なからずの家計がケインズ型消費関数に従っている、あるいはライフサイクル恒常所得仮説に従わない家計があることを意味している[7]。第 3 式の特定化が前 2 式に比べて優れているのは系列相関の検定である Durbin-Watson 統計量が 2 に近いことで、系列無相関になっていることである。決定係数が低いのは、前 2 式がレベル推定式であったのに対してこの式は成長率推定式であるためにフィットが落ちたからであるが、モデル特定化としては間違っていない。

これまでの消費関数の推定はマクロ時系列モデルを用いたものであり、ミ

[7] マクロ計量経済学でできることはここまでである。実際には、家計消費のうち、何家計がケインズ型消費をしており、何家計がライフサイクル恒常所得仮説型消費をしているのかを厳密に識別することが望ましいが、そのためにはミクロデータによって識別検定を行うしかない。

クロ計量経済学で扱うミクロデータに基づいたものではなかった。ミクロデータではどのような違いが生じるのかを直感するために図 3 を見てほしい。先ほどの図 2 とデータのバラつき具合が比較にならないほど大きいことがわかる。図 3 のデータはヴェトナムの家計支出調査に基づくものである。この調査では家計消費支出に関しては細かく調べられているが、家計構成員全員の所得および税を掌握することは難しく、可処分所得は分からない。そこで家計支出を合計した総家計支出と個別の消費支出の関係を見ていくことにする[8]。

図 3 左では OLS 推定の結果を推定値として入れてある。この推定式は具体的には次のようなものである。

$$\ln hhex12m = 0.935 + 0.573 \ln hh\exp 1$$

ここで $\ln hhex12m$ は家計医療費支出の対数、$\ln hh\exp 1$ は総家計消費支出である。すなわち、平均家計の医療費への消費性向は 0.57 程度である。しかし、データがこれだけバラついていれば平均からの乖離である OLS 推定の結果だけを見ていても多様な家計行動を掌握したことにはならないだろう。そこで図 3 右では区分回帰の結果を推定値として挿入しておいた。ここで用いた区分は家計医療費支出を下から 10% 、50% 、90% に分けてその区分に入る確率に基づいてパラメータを計算したものである。下から 10% の区分回帰の総家計消費支出への係数は 0.151（t 値 2.74）と極めて低い。50% の中位値では 0.621(t 値 16.00) と平均値よりも高くなり、90% では 0.80(t 値 15.47) と極めて高くなっている。この結果は、家計医療費支出の基準をどこに置くかによって推定結果がかなり違ってくることを意味しており、相当の多様性があることを示唆している。また、この家計医療費支出と総消費支出の関係は非線形である可能性もある。ここではこれ以上、この問題には踏み込まないが、ミクロデータの世界に入ると一気にモデルが複雑化し、経済主体の多様性に対応していかなければならなくなることがおわかりいただけたのではないかと思う。

4 おわりに

本章では計量経済学で最も基本的な線形モデルの推計方法や統計診断に関して説明を行った。また、金融マクロ時系列データとミクロデータではその性質も集計の仕方も違うので同じようなアプローチをとることが出来ないということも明らかにしたつもりである。

経済学では一般にモデルの経済的意味や特定化に関心があり、統計的にその仮説を検定することが計量経済学の役割だと理解している人が多いと思う

[8]家計が貯蓄をしないとすれば、総家計支出が可処分所得であると考えることが可能である。もちろん中には貯蓄をしている家計もあるだろうが、可処分所得そのものがデータとして手に入らない場合には、総家計支出を可処分所得の近似として扱っていいだろう。

が、実際には逆説的になるが、一般にモデルで説明できない誤差の振る舞いを細かく調べることで、そのモデルが適切であるかどうかを判断しているのである。すなわち、経済モデルの主要な説明変数で説明されていない部分が、純粋な誤差項で、不均一分散や系列相関を持たず、特定の変数と相関を持つことが無ければモデルとして棄却はしないが、これらの問題が一つでもあれば、何らかのモデルの改善を行う必要があり、また、どのように改善すれば良いかを検討しているのである。

ミクロ計量経済学では大量のサンプルを用いた分析を行うことが多く、サンプルに含まれる経済主体の多様性も増してくるので、単純なモデルで説明できる範囲は極めて限られたものにならざるを得ない。逆に言えば、計量経済学者は、その多様性を適切に整理し、分解することで、問題の本質を抽出していく作業に関与することが仕事なのである。

5 STATA コード

```
/**Production Function**/

use "cobb.dta", clear
/**data generation**/
gen k=exp(lnk)
gen l=exp(lnl)
gen y=exp(lny)
gen pery=y/l
gen perk=k/l

/*Ordinary Least Squares:OLS 推定*/
reg lny lnk lnl
test lnk+ lnl=1
hettest
ovtest
estimates store olssual
reg lny lnk lnl, robust
test lnk+ lnl=1
ovtest
estimates store olsrobust

/*Weighted Least Squares:WLS 推定*/
gen abslnk=abs(lnk)
```

```
reg lny lnk lnl [aweight=1/abslnk],robust/*資本ストックをウェイトとする*/
test lnk+ lnl=1
ovtest
estimates store wlsrobust
gen abslnl=abs(lnl)
reg lny lnk lnl [aweight=1/abslnl],robust/*労働者をウェイトとする*/
test lnk+ lnl=1
ovtest
/*Geleralized Least Squares:GLS 推定*/
gen lnksq=lnk*lnk
reg lny lnk lnl [aweight=1/lnksq], robust/*資本ストックをウェイトとする*/
predict lnyhat
test lnk+ lnl=1
ovtest
estimates store glsrobust
gen lnlsq=lnl*lnl
reg lny lnk lnl [aweight=1/lnlsq], robust/*労働者をウェイトとする*/
test lnk+ lnl=1
ovtest
/*Generalized Linear Models: GLM 推定*/
/*Maximum Likelihood Method*/
glm lny lnk lnl, family(gaussian) link(identity)
glm lny lnk lnl, family(gaussian) link(identity) robust
estimates store glmrobust

/*table*/
estimates table olsrobust wlsrobust glsrobust, se stats(N r2) b(%7.3f) keep(lnk lnl _cons)

/*Graphics　図 1*/
twoway (scatter pery perk)(fpfit pery perk), /*
*/ytitle (Percapita Output) /*
*/xtitle(Percapita Capital)/*
*/legend(label(1 "Actual Percapita Output") label(2 "Predicted Percapita Output"))
graph save "Production.gph", replace

/**Time series consumption function **/
```

```
use "consump.dta", clear

/**time series setting**/
tsset year

/**regression**/
/*level regression*/
reg rcons i3 inf rdisp
hettest
ovtest
dwstat
reg c y
predict chat
hettest
ovtest
dwstat

/*log linear regression*/
reg lc ly i3
predict lchat
hettest
ovtest
dwstat
/*dynamic linear regression*/
reg gc gy gc_1 gc_2 gy_1 gy_2 r3 r3_1 r3_2
hettest
ovtest
dwstat

reg gc gy
predict gcgy
hettest
ovtest
dwstat

/*graph*/
twoway (scatter lc ly)(line lchat ly)
graph save "consumption1.gph", replace /*図2右*/

twoway (scatter c y)(line chat y)
graph save "consumption0.gph", replace  /*図2左*/
```

```
graph combine consumption0.gph consumption1.gph
graph save "TSconsumption.gph", replace /*図 2 */

/**Vietnam Living Standard Survey Data**/
use "vietnam_ex1.dta", clear

reg lhhex12m lhhexp1
predict pols
reg lhhex12m lhhexp1, robust

twoway (scatter lhhex12m lhhexp1)(line pols lhhexp1), ytitle(Log Household Total Expenditure) xtitle(Log Household Medical Expenditure)/*
*/ legend(pos(11) ring(0) col(1)) legend(size(small)) /*
*/ legend( label(1 "Actual Data") label(2 "Mean"))
graph save "consumption2.gph", replace   /*図 3 左*/

* Bootstrap standard errors for OLS
set seed 10101
* bs "reg lnmed lntotal" "_b[lntotal]", reps(100)
* (1) Quantile and median regression for quantiles 0.1, 0.5 and 0.9
* Save prediction to construct Figure 4.2.
qreg lhhex12m lhhexp1, quant(.10)
predict pqreg10
qreg lhhex12m lhhexp1, quant(.5)
predict pqreg50
qreg lhhex12m lhhexp1, quant(.90)
predict pqreg90

graph twoway (scatter lhhex12m lhhexp1) (lfit pqreg90 lhhexp1) /*
*/ (lfit pqreg50 lhhexp1) (lfit pqreg10 lhhexp1), /*
*/ xtitle("Log Household Medical Expenditure") /*
*/ ytitle("Log Household Total Expenditure") /*
*/ legend(pos(11) ring(0) col(1)) legend(size(small)) /*
*/ legend( label(1 "Actual Data") label(2 "90th percentile") /*
*/ label(3 "Median") label(4 "10th percentile"))
graph save "consumption3.gph", replace
/*Cameron and Trivedi (2005, Figure 4.2 p.90) の図の再現、図 3 右*/

graph combine consumption2.gph consumption3.gph
graph save "CSconsumption.gph", replace   /*図 3 */
```

参考文献

[1] 北村行伸 (2005) 『パネルデータ分析』、岩波書店

[2] 畠中道雄（1991）『計量経済学の方法』、創文社

[3] Amemiya, Takeshi.(1985) *Advanced Econometrics*, Blackwell.

[4] Baum,Christopher F.(2006) *An Introduction to Modern Econometrics using Stata*, Stata Press.

[5] Breusch, T.S. and Pagan,A.R.(1979) "A Simple Test for Heteroskedasticity and Random Coefficient Variation", *Econometrica*, 47, pp.1287-1294.

[6] Cameron, A.C. and Trivedi, P.K.(2005) *Microeconometrics: Methods and Applications*, Cambridge University Press.

[7] Davidson, Russell and MacKinnon, James G.(2004) *Econometric Theory and Methods*, Oxford University Press.

[8] Koenker, Roger. (2005) *Quantile Regression*, Cambridge University Press.

[9] White, Hilbert.(1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48, pp.817-838,

[10] Winklemann, Rainer and Boes, Stefan.(2005) *Analysis of Microdata*, Springer.

[11] Wooldridge, Jeffrey. M.(2003) *Econometric Analysis of Cross Section and Panel Data*, The MIT Press

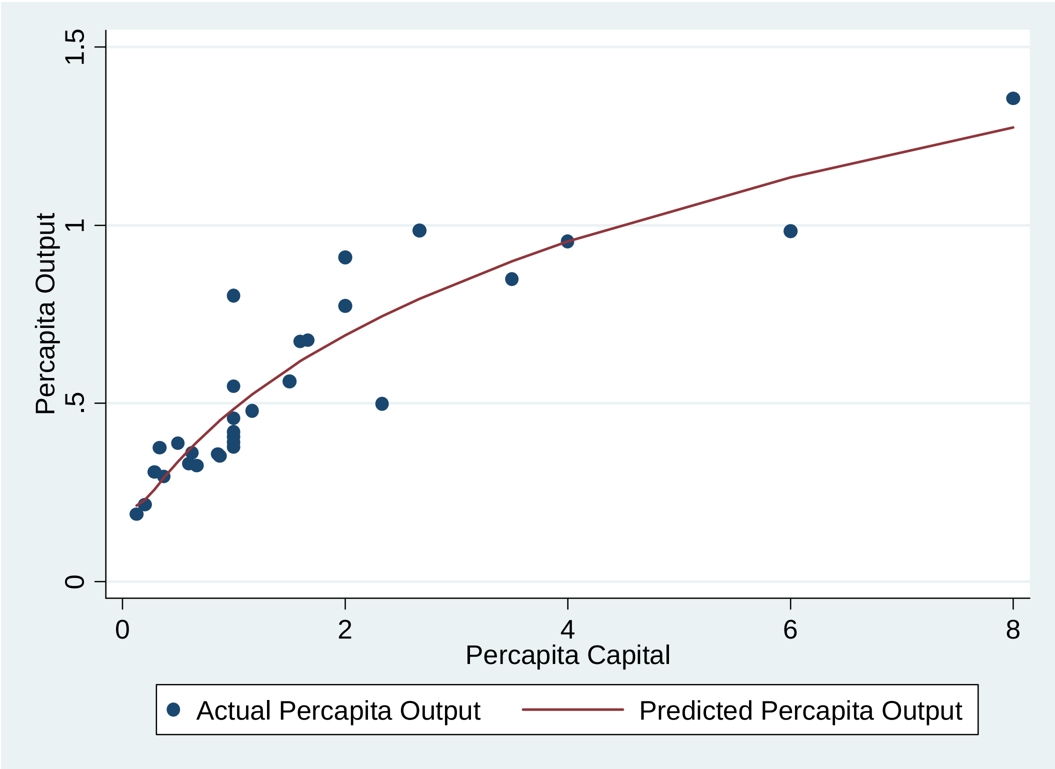
表1 コブ･ダグラス型生産関数の推定

被説明変数:lny	OLS		WLS		GLS		GLM	
	Coef.	Robust-t	Coef.	Robust-t	Coef.	Robust-t	Coef.	Robust-t
説明変数								
lnk	0.347	7.75	0.343	7.96	0.340	8.14	0.347	8.03
lnl	0.414	11.75	0.407	11.03	0.397	10.23	0.414	12.18
_cons	0.450	1.53	0.509	1.71	0.571	1.91	0.450	1.58
観察値	30		30		30		30	
F(2,27)	95.20		86.65		79.10			
R-squared	0.875		0.88		0.882			
Residual df							27	
Scale prameter							0.240	
(1/df) Deviance							0.240	
(1/df) Pearson							0.240	
AIC							-0.802	
BIC							-91.187	
test lnk+lnl=1	F(1,27)=16.76 Prob>F=0.000		F(1,27)=18.17 Prob>F=0.000		F(1,27)=19.63 Prob>F=0.000			
Breusch-Pagan/Cook-Weisberg test for heteroskedasticity	chi2(1)=0.25 Prob>chi2=0.6188							
Ramsey RESET test	F(3,24)=1.48 Prob>F=0.246		F(3,24)=1.76 Prob>F=0.182		F(3,24)=2.10 Prob>F=0.127			

表2 時系列データによる消費関数の推定

被説明変数	c		lc		gc	
	Coef.	t	Coef.	t	Coef.	t
説明変数						
y	0.779	112.79				
ly			0.956	109.03		
i3			−0.002	−3.00		
gy					0.571	8.47
_cons	463.179	4.69	0.229	2.81	0.008	4.25
観察値	37		37		36	
R−squared	0.997		0.998		0.679	
Adj R−squared	0.997		0.998		0.669	
Root MSE	133.09		0.011		0.007	
Breusch−Pagan/Cook−Weisberg test for heteroskedasticity	chi2(1)=1.14 Prob>chi2=0.285		chi2(1)=0.36 Prob>chi2=0.550		chi2(1)=1.23 Prob>chi2=0.267	
Ramsey RESET test	F(3,32)=10.02 Prob>F=0.000		F(3,31)=10.02 Prob>F=0.000		F(3,31)=3.67 Prob>F=0.023	
Durbin−Watson statistic	(2,37)=0.804		(3,37)=0.686		(2,36)=2.115	

図1　一人当たり生産量



Percapita Output
0
.5
1
1.5
0
2
4
6
8
Percapita Capital
Actual Percapita Output
Predicted Percapita Output

図２　一人当たり消費量と可処分所得の関係

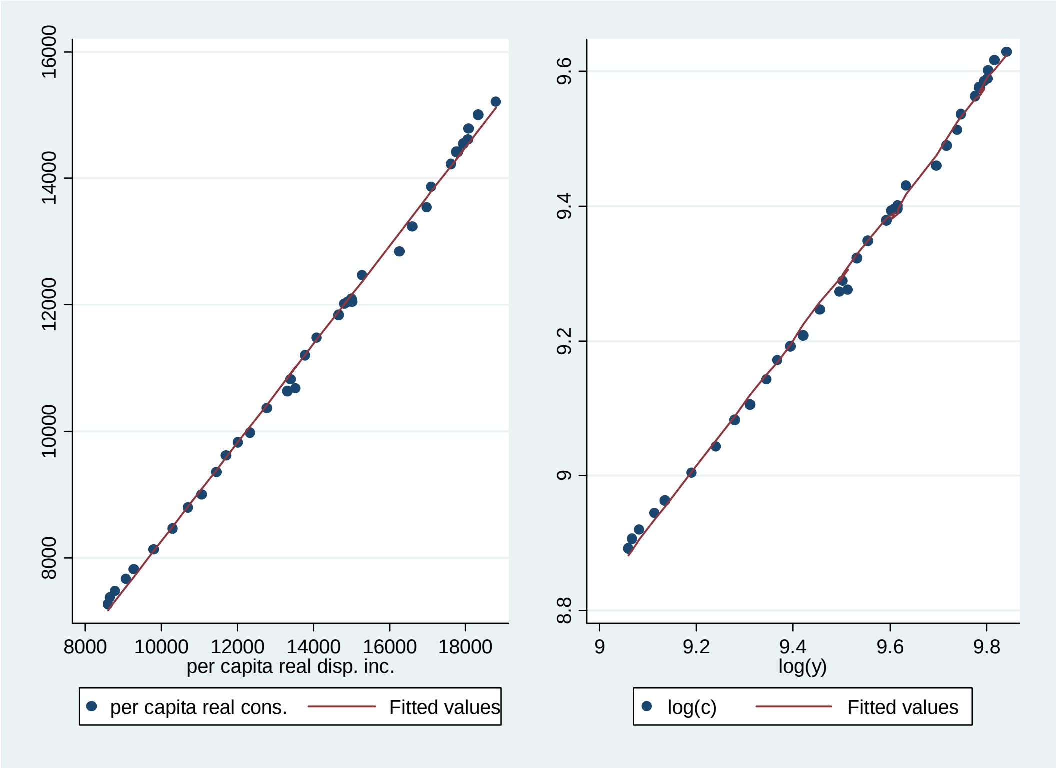


図3　家計医療費支出と総家計消費支出の関係