

経済学と反事実分析 接触篇

Economics and Counterfactual Analysis: A Contact

ill-identified

2020 年 1 月 25 日

1 概要

第 83 回 Tokyo.R のエントリ. 前回の戦後処理に時間を取られたので, 手抜き回.

前回にも言及した Igami (2018) の研究に関連して, **構造推定**についてももう少し詳しい話をしようと思っていたが, 残り時間の問題から**動学構造推定**は難しいと判断した. しかしアンケート至上主義なので経済学のネタは外せないから, 準備の簡単な**静学構造推定**の話を取りあえずやることにした. 静学構造推定は「動学」ではないので時間変化を考えないタイプの理論モデルを推定する一連のテクニックである. 今回の話は比較的簡単かつ昔から知られている話だから大学の経済学部でも授業で取り上げられることもあろう. 知ってる人も多いかもしれない.

具体的には, 生産関数を構造推定する OP 法 (Olley and Pakes, 1996) および LP 法 (Levinsohn and Petrin, 2003) を R でやる. 実はこれらは既に計算プログラムが `estprod` パッケージによって提供されているが, 使って終わりにするのではなく仕組みを解説しつつ使う.

2 イントロダクション

2.1 構造推定とは

では構造推定とは何だろうか. 構造推定で検索すると, 名前が似ているが違うものがいくつか出てくる.

■**構造方程式モデリング** 例えば構造方程式モデル (SEM) というものがある. これは変数間の関係構造を同時線形方程式で表したもので, 因果推論のフレームワークの原型ともいえるものだが, **構造推定とは違う**.

■**構造時系列モデル** 構造ベクトル自己回帰モデル (Structural VAR) とか, **ベイズ構造時系列モデル (BSTS)** というものがある.

通常の VAR は複数の変数の時系列変化を扱うモデルで, それぞれの変数が過去の値にのみ依存して決まる回帰モデルなのに対して, 構造 VAR は現在の別の変数にも依存する, **同時決定的な時系列モデルである**^{*1}. ペ

^{*1} 簡単な定義と特徴は<https://qiita.com/saltcooky/items/2d0119ea4a10bab6cff2>の最後のほうに書いてある. より詳しい話を知りたいければ, 沖本 (2010) が言及している. インターネットの日本語圏で出回っている情報の多くはこのあたりからだと思うが, もっと多変量時系列モデルを知りたいなら Lütkepohl (2007) の教科書がある (この本は約 700 ページの間, タイトル通りマジで多変量時系列モデルのことしか書いてない). 時系列モデルに関する本は北川 (2005) の教科書を始め和書でも多く出ているが, 経済学や

イズ構造時系列モデルの「構造時系列」は Harvey のいう構造時系列モデルのことを指す (Scott and Varian, 2014). これは 89 年とかなり古い教科書なので私も読んだことがないが, Lütkepohl (2007, p. 618) によればローカルトレンドモデルを基本とした時系列モデルらしいので, 多変量時系列モデルというわけでもなく, 状態空間モデルに属する. これらも**構造推定とは関係ない**.

■**構造型モデル** いちおう過去の Tokyo.R の発表ネタとかぶらないか確認したところ, こんなものを見つけた.

R で学ぶ『**構造型モデル de 倒産確率推定**』

「構造型モデル」という名前も構造推定と似ている. いちおう元になった Merton (1974) の論文も確認した. これも**構造推定ではない**.

構造推定の定義について, 多くの人は Reiss and Wolak (2007) の紹介を挙げるが, この本はとにかく分厚くて高いので私は持っていない. 前回で触れたように, 観察されたデータにそのまま回帰式や他の関数を当てはめても因果関係を知ることはできない. そこで, いくつかの要素が必要になる. そこで構造推定は仮定によって構造を特定しようとした.

北村 (2016) は一般的な型はないとしつつ, その解説を要約すると 3 つの要件となる.

1. 経済理論モデルであること
2. 確率モデルであること
3. データに基づいて推定されていること

(1) はいわゆるルーカス批判 (Lucas, 1976) を受けて, 彼の批判を克服できるような経済理論に基づいたモデルづくりを意識するようになった. (2) はちょっとわかりづらい. 回帰分析だって正規分布を暗に仮定しているから確率モデルになる. しかしここでは, 観察できないものと, できるものの不確実性を考慮しているという意味である. 例えば測定誤差や, 分析者には観測できないが, 経済活動の当事者は知っている情報というものがある. これらを無視した推定はしばしばバイアスを引き起こすが, 構造推定ではモデルにこのような仮定を取り込む. (3) は, 完全な架空の話ではなく, なるべく現実と整合するために観測されたデータに基づいてパラメータを決定するということである.

よって, 最初に挙げた「構造方程式モデリング」や「Merton の構造型モデル」は経済理論に基づいた構造を表現していないため構造推定ではない. ただし構造時系列モデルは, 実は 80 年代にマクロ経済モデルとして利用された歴史がある. 経済理論に基づかないという批判は当時からあったものの, 単に予測するだけなら問題ないということで使われていた. しかし, 「どうなるか」と「なぜそうなるのか」のどちらを求めているかと言えば前者であり, 原因を分析することはできない^{*2}.

一般に定まった方法はないが, かといって完全な無手勝流ではない. 大学院のコアコースで教えるモデルはだいたい決まっており, 多くの研究も先行研究のフレームワークを応用していることが多い.

2.2 因果推論と何が違うのか

以前にも紹介したような RCT, 準実験 (操作変数法, 傾向スコア法, DID, RDD) など, 「反事実的な因果推

数理ファイナンスの観点から書いたもの以外では, 構造 VAR について言及したものは私の知る範囲では見られない. 構造 VAR という考え方自体がパラメータの識別を前提としているからだろう. マクロ経済学における時系列モデルの利用を知りたいのなら, 渡部 (2016), 中島・渡部 (2012) が参考になる. 本格的な教科書は専門外なので知らない.

^{*2} そのため, VAR でも因果関係を定義できるということで Granger 因果性が注目を集めた. つまり, 以前の投稿で問題にした, 予測か因果かという二択は, 経済学の発展のなかでもしばしば議論されてきた問題だということがわかる.

論」と呼ばれる。しかしこれらは経済学の理論仮説に基づく因果関係ではなく、ランダム化と比較によって「平均処置効果」を求めるフレームワークである^{*3}。一方で構造推定は実際のデータと理論モデルを融合させた結果なので、本来の意味で現実には起こり得なかったこと（反事実）を推論するフレームワークである。

具体例を挙げよう。今回は経済学における生産関数の推定の例を紹介する。原材料を投入し、財を生産する、その変換ルールを関数に表したものを生産関数という。例えば、資本 (K apital^{*4}) と労働力 (L abor force), Y は出力、つまり生産量なので、生産関数を以下のように表せる

$$Y = F(K, L)$$

この生産関数の形状が分かれば、

- 実際のデータとの残差から、数値に表れない生産性 (TFP) を計測する (残差分析)
- もし設備投資を増やしたら (K を増加させたら) どれくらい生産量が増えるか
- もしパラメータが変化したら、生産力はどう変化するか (比較静学分析)

と言ったことがわかる。経済学部出身者ならば、ミクロ経済学の授業で聞いた覚えがあるだろう。しかし、データから推定した係数が「もし変わったら」というのは、本質的に現実に起こっていない事象について話していることになる。同等のことを因果推論でやるとすれば、両方のケースについて実際に観測したデータに基づいて差を比較しなければならない。つまり、単にデータを近似しただけの関数では、これらのことを言う根拠として弱い。

経済学では RCT や準実験アプローチは誘導形 (reduced form) 推定^{*5}とも呼ばれ、構造推定派と誘導形推定派は数年前までお互い相手のアプローチの問題点を指摘し激しく論争していた。その結果、現在では現在はそのそれぞれのアプローチでできることの限界がはっきりしてきたため、論争というほどのことは無い^{*6}。

両者の使い分けとしては、構造推定は経済理論モデルというある意味「架空のもの」に基づいて未来に起こりうることについての分析もできるが、因果推論は実態としては実験計画法と本質的に同じで統計学の理論に基づいて観察できたデータから事後評価するアプローチである。よってしばしば、構造推定は反事実 (反実仮想) シミュレーション、因果推論はプログラム (政策) 評価とか介入 (intervention) といった呼び分けられる^{*7}。そして構造推定は理論仮説に強く依存したフレームワークであるので、仮説が妥当なのかはよく注意が必要である。ただし、かといって因果推論が仮定に依存しないというのも完全な誤りである。前回でも言及したが、因果推論であっても必ず仮定が存在するので、それを無視して計算した平均値の差が因果効果を意味するとは限らない。

^{*3} 正確には少し違うが、重要なのはサンプルの平均的な差を見るという特徴が共通している

^{*4} ドイツ語でのスペル。C にしてしまうと消費, Consumption と重複するからだと思われる。また、正確には「資本ストック」であり、設備投資や固定資産などから計算する

^{*5} 構造推定に使うような、経済現象の構造を表したモデルを、平均の差の比較という因果推論のフレームワークで推定できるように単純な形に変形することからこう呼ばれる。この式変形のため、根本的な構造パラメータを特定しづらいという問題がある。

^{*6} 歴史的経緯は北村 (2016), 中嶋 (2016) などが少しだけ解説している。計量経済学の観点では統計モデルとして Hayashi (2000, Ch. 8) が言及している。また, Cameron and Trivedi (2005, Ch. 2) は教科書の最序盤で構造モデルと誘導モデルの特徴付けを行うという独特の構成になっている。

^{*7} 後者の因果推論フレームワークも、もともとは counterfactual causal model と呼ばれていたが、(おそらくはこの性質の違いのため) 少なくとも経済学の分野では最近あまり強調されていない気がする。

3 今回取り組む問題

動学構造推定はテクニカルでなんかすごいことをやってるという印象を簡単に与えられる。しかし、今回は時間がないのでもう少し簡単にできるほうをやる。前回話題にしたのは「動学的」構造推定であり、時間変化を取り入れたモデルを扱うジャンルである。一方で、今回は考慮しない「静学的」構造推定の具体例を紹介する。といっても昔からある研究なので、すでに日本語の講義スライドもネット上に転がっていたりするし、もしかしたら発表者の中にも授業でやったことがあると言う人がいるかもしれない。

今回は特に、生産関数の推定の問題を例にする。

3.1 生産関数

生産関数をどういう形にするかはいろいろだが、よく使われるものの1つにコブ = ダグラス型生産関数がある^{*8}。特に、一企業の

$$Y = f(K, L) := \alpha K^{\beta_K} L^{\beta_L}$$

という指数関数で表す^{*9}。 α, β_K, β_L はパラメータである。 $\beta_K = \beta_L = 0.5$ の場合、図 1 のような形状になる。

3.2 なぜ生産関数を知りたいのか

反事実分析 (比較静学)

パラメータの増加に対してどう変わるか

4 どう推定するか

経済学部で計量経済学の講義では、回帰分析で生産関数の推定をやるところも多いと思う。 Y を被説明変数、 K, L を説明変数としても、線形式ではない。しかし、 Y, K, L がいずれもゼロ以下になりえないことに着目すると、対数を取って以下のようなになる (いわゆる対数線形回帰モデル)。

$$\ln Y = \ln \alpha + \beta_K \ln K + \beta_L \ln L$$

となる。つまり、おなじみの最小二乗法 (OLS) で計算できる。観測できなかった α も、切片パラメータとして推定できることになる。以降はそれぞれの対数を y, β_A, k, l に置き換えて、

$$y = \beta_A + \beta_K k + \beta_L l + \varepsilon \tag{1}$$

^{*8} 他にも色々あるが、例えば物理学のように法則がかなり明確になっている分野を勉強した人間からはおかしいかのように聞こえるかもしれない。しかし、経済学ではむしろ、変に特定化すれば特定化の誤りを免れない。Cobb-Douglas 型関数には計算上の利便性もあるが、ロバストな関数を用いるたほうがよいと考えられる。

^{*9} 生産関数のより詳しい数学的性質はミクロ経済学の教科書に書かれている。例えば、特に有名なのは西村 (1990)、『ミクロ経済分析』(Varian, 1992)、Mas-Colell et al. (1995) などである。それ以降にも評価の高いミクロ経済学の和書がいくつかでていますが、私はあまり読んでいないのでなんとも言えない。

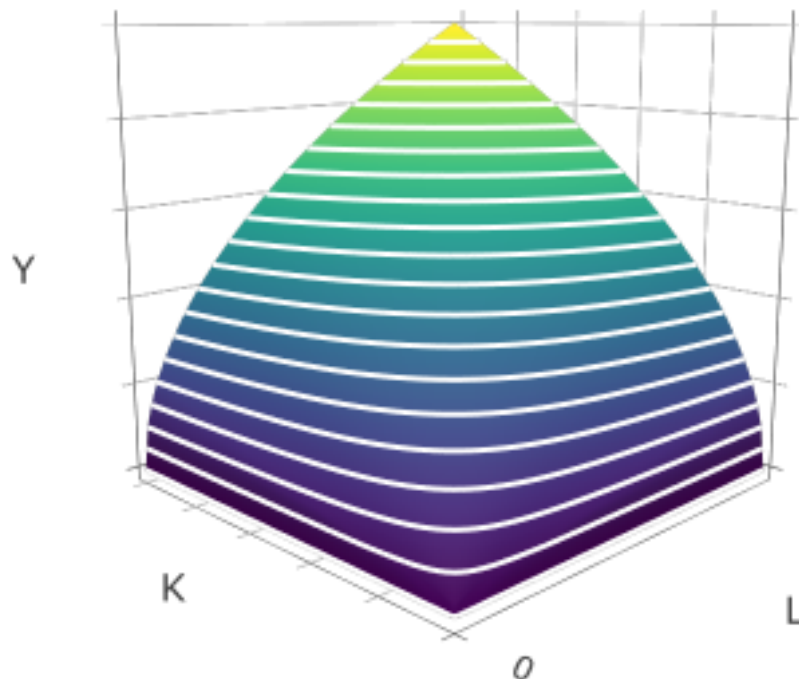


図1 コブ＝ダグラス型生産関数の例

という線形回帰モデルを考える。しかし以前書いたように、経済学の文脈では非線形だろうが線形式だろうがデータに回帰を当てはめただけでは用をなさない。経済理論に即して、発生しうる問題を排除した方法で推定しなければならない。つまり、現在では OLS では生産関数のパラメータを**推定するには不十分**と考えられている。

発表時間の問題から、スライドでは大幅に省略しているが、ここでは冗長気味にパネルデータ分析の発展の経緯も合わせて説明していく^{*10}。

現在の経済学の観点からすると、(1)をそのまま OLS で推定するのをためらう大きな理由は以下の 3 つである。

1. 企業特有の効果が存在する
2. 需要要因を考慮していない
3. セレクションバイアス

4.1 パネルデータ分析と観測されない効果

「企業特有の効果」とは、特定の企業だけが「特殊な技術を持っている」「優秀な人材が集まっている」など、単純な量だけでは捉えきれない情報を意味する。つまり、企業ごとに異なる ω_i が生産量に影響していると考える。

$$y_i = \beta_A + \beta_K k_i + \beta_L l_i + \omega_i + \varepsilon_i$$

^{*10} こういった初歩的な話は、私なんかよりよほど深く理解した大学の先生がいくつも教科書を書いており、正確さや詳しさをそちらのほうが遥かに良い。参考文献としては、一般公開されている奥井 (2016) のスライド、Wooldridge (2010, Ch. 10-11) の教科書などがある。私はほぼ Wooldridge の教科書で勉強したため、日本語の教科書に詳しくない……

企業 ID	時期	生産量	資本	労働力
001	2019Q1	100	10	10
001	2019Q2	110	11	10
001	2019Q3	90	12	12
001	2019Q4	120	14	13
002	2019Q1	80	5	5
002	2019Q2	90	6	5

表1 パネルデータのイメージ

ω_i は観測できないので、このまま (y_i, k_i, l_i) のデータで回帰すると、 ω_i は定数項 β_0 または誤差項 ε に吸収される。 ω_i が k_i, l_i と全く相関しない変数ならば、 β_A と ω_i を識別できないものの、少なくとも β_K, β_L の推定はできる。

このような問題をどう解決すれば良いだろうか？ まずは問題 1 だけを考えよう。 ω_i は観測できないが、では推定できないだろうか？ つまり、同一の企業の情報を何度も観測しすれば、そこから推定できないか、というアイデアである。これがパネルデータ分析のスタート地点になる。パネルデータとは、同一個体を追跡して複数時点で観測したデータである (表 1)。よって、企業 $i = 1, \dots, N$ ごとに、 $t = 1, \dots, T$ の観測がある。そこで以降は回帰モデルを以下のように書く。

$$y_{i,t} = \beta_A + \beta_K k_{i,t} + \beta_L l_{i,t} + \omega_i + \varepsilon_{i,t} \quad (2)$$

ここではひとまず、 ω_i が企業 i ごとに異なるが、時間 t に対しては変化しないと考えておこう。また、問題 1 だけを考えるにあたって、 ω_i は $k_{i,t}, l_{i,t}$ と相関しないとしておこう。このような ω_i を観測されない個別効果 (unobserved individual effect) という。

さらに追加の仮定として、 ω_i が定数である場合、ダミー変数を作成することで OLS であるかのように ω_i を推定できる^{*11}。これを LSDV 推定量という。また別の解き方としては、(2) を企業ごとに平均を取った場合を考える。

$$\bar{y}_i = \beta_A + \beta_K \bar{k}_i + \beta_L \bar{l}_i + \bar{\omega}_i + \varepsilon_{i,t} \quad (3)$$

このとき、 ω_i は企業ごとに一定なので、 $\bar{\omega}_i = \omega_i$ である。よって、(2) から (3) を引いてから OLS で計算することでも同じ推定結果が得られる。こちらの方法は固定効果 (FE; fixed effect) 推定量またはグループ内 (within) 推定量と呼ばれる^{*12}。

一方で、 ω_i を確率変数として扱うことも考えられる。その場合は変量効果 (RE; random effect) 推定量と呼ばれる。この場合、数式上は分散不均一な回帰モデルと同一のため、実質的に一般化最小二乗法 (GLS) と同じである^{*13}。

^{*11} ただし、一致するのは期待値のみで、標準誤差の計算方法は変わる。仮説検定・区間推定をするならば clustering-robust な標準誤差と呼ばれるものを使う。

^{*12} 逆に観測されない効果が時間ごとに異なるが、企業間では一定の場合も同じようにできるだろう。その場合はグループ間 (between) 推定量と呼ばれる。

^{*13} 久保 (2012) は生物学の文脈で異質効果のある場合の解法として一般化線形混合モデル (GLMM) を紹介している。これは観測されない効果の問題とよく似ている。生物の実験データではおそらく変数との相関がないため、定数項のバイアスと、当てはまりの良さだけが問題となるのだろう。

4.2 動学パネルデータ分析

このように、観察できない要因があってもパネルデータがあれば理論上は推定が可能になる。しかし、Mairesse (1990)^{*14}はこれまで紹介した方法とほぼ同じやり方で企業の生産関数を推定したところ、観測されない効果を識別しようと複雑に差分を取れば取るほど、係数の推定値が小さくなることを発見した。彼はこの原因の仮説をいくつか挙げている。

1. 内生性の問題。個別効果 ω_i は物理的なものではなく、環境、経営者の手腕といったものに由来するのなら、他の説明変数 k, l の決定に影響している。
2. 企業的意思決定のラグ。経営者が誤差項として現れる短期的な変化、つまり直前の $\varepsilon_{i,t-1}$ に対応して $k_{i,t}, l_{i,t}$ を決定しているなら、誤差項と説明変数が相関しており、単純な差分では推定できない。
3. 同時決定の問題。
4. 測定誤差。各変数に測定誤差があれば、差を取ることで打ち消される。

これらの問題への対応として考えられるのが、**操作変数**である。

(2)に、被説明変数 $y_{i,t}$ の自己回帰項を追加する。

$$y_{i,t} = \beta_A + \beta_K k_{i,t} + \beta_L l_{i,t} + \rho y_{i,t-1} + \varepsilon_{i,t} \quad (4)$$

こういう自己回帰項を含むモデルを扱う分野はパネルデータ分析のなかでも特に**動学パネルデータ分析**と呼ばれる。このあたりの問題も奥井 (2016) や Wooldridge (2010, Ch. 11) で丁寧に説明されている^{*15}ので冗長だが、一応簡単に述べておく。

前節のアイデアを流用すれば、Within 推定量の応用で時間平均を取ることが考えられる。 $\bar{y}_i = (T-1)^{-1} \sum_{t=2}^T y_{i,t}$ しかし、パネルデータでは N が大きい T は小さいことが多いため、一致推定量とみなせない。一方で、**1階差分 (FD)** 推定量というものがある。これはある意味 **DID 推定量**と同じである。差分演算子を $\Delta x_t := x_t - x_{t-1}$ というふうに定義する。(4)の両辺から $y_{i,t-1}$ を引く。

$$\Delta y_{i,t} = \beta_K \Delta k_{i,t} + \beta_L \Delta l_{i,t} + \rho \Delta y_{i,t-1} + \Delta \varepsilon_{i,t}$$

しかし、 $\Delta y_{i,t-1} = y_{i,t} - y_{i,t-1}$ から、 $\Delta \varepsilon_{i,t} = \varepsilon_{i,t} - \varepsilon_{i,t-1}$ と相関が存在する。そこで、 $\varepsilon_{i,t-1}$ より過去の $y_{i,t-2}, \dots$ を**操作変数**として使う方法が提案されている。それが Anderson-Hsiao (1981; 1982) モデルや、Arellano-Bond (1991) のシステム GMM 推定法である^{*16}。しかし、これらは実際に操作変数として意味のある変数なのかということを深く議論せずに $\{y_{i,t-2}, \dots\}$ を操作変数に利用している。その場合、上記で挙げたような意思決定が存在するかどうかを検証できないし、推定量じたいにも過剰識別や弱相関変数の問題が発生する。

^{*14} 古い本なので大学図書館などでないと閲覧が難しそうだ。私も手元にないので学生の頃にかいたメモを元にも書いている。また、川口先生の講義ノートではこのあたりの問題に関する参考文献として Griliches and Mairesse (1999) の論文を挙げている。タイトルからしてもこちらのほうがカバー範囲の広いサーベイのようなので、こちらを参照したほうがいいのかもかもしれない。

^{*15} GMM も以前 **ブログ**に定義を書いたが、詳しいことは Hayashi (2000)、また、他の教科書として動学パネルデータを重点的に解説した千木良他 (2011) などがある。Baltagi の教科書も授業でおすすめされていた気がするが私は読んでいないのでどこまで書いてあるか知らない。どうしても本を買いたくない/買えない人は奥井亮『**動学的パネルデータモデル**』の講義スライドを見ると良い。

^{*16} さらに非線形方程式に拡張した Blundell and Bond (1998) の推定法もある。

4.3 企業内部の意思決定を考える

しかし、経済学の理論では、観察される経済データは市場が需要と供給のバランスを取った結果である。これが2つ目の「需要要因を考慮していない」という話になる。企業が需要を全く考えずに事業計画を立てるとは考えにくいので、この ω_i が間接的に k_i, l_i に影響している可能性は十分ある。また、企業は新規立ち上げだったり倒産したりする。現実のパネルデータはしばしば欠落がある。そのようなデータでそのまま推定しようとしたらどうだろうか。

このようなパネルデータでの内生性に対処する推定方法に関する研究はいくつもあり、特によく出てくるものとして FEIV モデル¹⁷がある。

さらに、[Olley and Pakes \(1996\)](#), [Levinsohn and Petrin \(2003\)](#) はこれらの動学パネルデータ分析の方法から一歩進んで、企業の意思決定の構造をモデル化した生産関数を考え、操作変数を使わずに内生性の問題に対処した方法を提案している（一般には OP の方法は **コントロール関数アプローチ**と呼ばれる）¹⁷。

そこで今回は [Olley and Pakes \(1996\)](#) のフレームワークを詳しく説明する。なお、以降は(2)式をもとに説明するが、各企業で同じことが言えるため i を省略して書く。OP では、観察されない効果 ω_t は企業の経営者は知ることができるが、データとして現れないため分析者には観測できない、企業固有の生産能力を表す成分である。 ω_t は他の変数の影響を受けない（外生的）が、自身の過去の値からは影響を受けると仮定する。つまり1次の **マルコフ過程**であるとする。

$$\omega_t \sim P(\omega | \omega_{t-1})$$

t 期の始めに経営者は ω_t を観測して、これをもとに今期の経営計画を決定、つまり、純キャッシュフローの割引現在価値を最大化できるように k_t, l_t を決めていくと仮定する。

そこで、動学的なモデルを想定する。企業の t 時点現在での収益を $\pi_t(\omega_t, k_t, l_t)$ とし、投資 inv_t に対する費用を $c(inv_t)$ とする。経営者が最適化する現在の **価値関数 (value function)** を $V_t(\omega_t, k_t, l_t)$ とすると¹⁸、動学モデルでは単に現在の価値関数を最大化することが最適化に繋がるとは限らない。経営者は、これらを把握した上で、割引現在価値を最大化するように行動する。つまり、以下のような V^* の達成を目指す。

$$V^* = \max_{\{k_t, l_t\}} \sum_{t=0}^{\infty} \delta V_t(\dots)$$

これは無限和なのでそのまま計算するのは難しいが、一定の条件下では上記の最適化問題の解が以下のような **ベルマン方程式 (Bellman equation)** の解と同じであることが分かっている¹⁹。

¹⁷ 日本語では、たとえば [北村他 \(2009\)](#) が理論的根拠を詳しく解説している。彼らの解説で使われているのは本来のものに加えて原材料費や純資産を考慮したモデルだが、本質的な部分は変わらない。統計モデルとしての説明なら彼らのものでも十分だが、今回は経済モデルとの関係を強調したいため、あえて Olley らの論文をなぞって冗長に書いておく。

¹⁸ Olley らの当初の仮定では、企業年齢も生産性に影響があると考えていたが、実際に推定してみるとほとんど影響がないことがわかった。モデルの特徴を説明する上でも企業年齢はあまり関与しないため、ここでは省略している。

¹⁹ これを **ベルマンの最適性原理**という。しかしこれ以降の動学最適化問題の数理的に厳密な話をしていくにはキリがないので、適当な参考文献に投げることにする。ただし私は動学最適化を純粋に数理的側面で説明した教科書を読んだことはなく、経済学の文脈で主に勉強しているので、挙げる参考文献も必然的にそういうものになる。日本語文献でかつ無料で一般公開されているものとしては、マクロ経済学での応用の観点で書かれた [工藤 \(2007\)](#), [阿部 \(2017\)](#), [蓮見 \(2019\)](#) の講義ノートがある。教科書では A.C. Chang の『動学的最適化の基礎』の邦訳は比較的簡単に解説しているが誤植が多いことで有名である。西村清彦の『経済学のための最適化理論入門』という教科書もあるが、どこまでカバーしていたかの記憶が定かではない。もっとも詳しいのはやはり [Stokey et al. \(1989\)](#), [Ljungqvist and Sargent \(2018\)](#) の教科書あたりか？単に「ベルマン方程式」でググると強化学習に関連するものとして言

$$V_t(\omega_t, k_t, l_t) = \max_{k_t, l_t} \left\{ \Phi, \sup_{i_t \geq 0} \pi_t(\omega_t, k_t, l_t) - c(inv_t) + \delta E[V_{t+1}(\dots)] \right\} \quad (5)$$

ベルマン方程式が意味するところは、 $t, t+1$ という 2 時点間の関係さえ分かれば上記の無限和の問題も解けるということである。

ただし、このモデルでは 2 つの仮定を負っていることに注意する。1 つは、(5) には、 Φ という選択肢があることだ。これは退出、つまり**工場の閉鎖や事業を売却**することで得られるキャッシュフローである。よって、このモデルでは採算が合わないと判断した経営者が事業撤退を選択する可能性も考慮している。また、動学的な決定の必要があるのは k_t のみで、 l_t は每期独立して決定されるように**仮定している**。

では、 k_t の動学的な意思決定がどのようになされるかを考えていく。資本 k_t は、前期のストック k_{t-1} と、減価償却率 δ 、そして新規の投資額 inv_{t-1} の和で表現できる。

$$k_t = inv_{t-1} + (1 - \delta)k_{t-1}$$

企業の退出ルールと投資関数を以下のように定義できる。

$$x_t = \begin{cases} 1 & \text{if } \omega_t \geq \underline{\omega}_t(k_t) \\ 0 & \text{otherwise} \end{cases}$$

$$inv_t = inv_t(\omega_t, k_t) \quad (6)$$

ここから、企業が退出する確率は

$$\begin{aligned} p_t &:= P(x_t = 1 \mid \underline{\omega}_t(k_t)) \\ &= P(x_t = 1 \mid k_t, \omega_t, \underline{\omega}_t) \end{aligned} \quad (7)$$

のように $k_t, \omega_t, \underline{\omega}_t$ 条件に依存することを覚えておこう。

利潤関数 π_t が k_t に対して減少関数なら、価値関数にとっては増加関数、そして $\underline{\omega}_t(\cdot)$ は減少関数になるはずだ。資本ストックの多い企業は現在の生産性とはあまり関係なく、将来に比較的大きな収益が見込まれるから、このしきい値 $\underline{\omega}_t(\cdot)$ は比較的低くなるはずだ。よって、この退出ルールから ω_t に $x_t, k_t, l_t, \omega_{t-1}$ を条件付けた ω_t の期待値

$$E[\omega_t \mid k_t, \omega_{t-1}, x_t = 1]$$

は k_t に対して減少関数であるとわかる (l_t は独立しているので ω_t に影響しない)。これが観測できない生産性 ω_t と資本が相関している理由付けになる。

ここまでで登場した変数間の関係をグラフィカルモデルで描くならば、以下のようになる。

及する記事がよく引っかかるが、これらはベルマン方程式というより分枝限定法の考え方の解説だったり、内容がメモ程度の断片的だったりするので今回の話の補足説明には向かない。理論モデルでは連続時間を想定する場合もあり、ベルマン方程式の代わりに Hamilton-Jacobi-Bellman 方程式を応用することもある。こちらは物理学でもよく使われるようだ。

4.4 OP モデルの回帰モデルとしての意味

経営者が(分析者には)観察されない効果 ω_t を参照して k_t, l_t を決めてしまうというのは、計量経済学の文脈でいうならこれは欠落変数バイアスであり、この仮定が正しければここまでで紹介したパネルデータの方法で推定してもバイアスが発生してしまう。さらに ω_t のマルコフ性を仮定しているため、もし ω_t が過去の値と相関しているならば、やはり内生性によるバイアスが発生する。また、OP モデルは経営者の撤退という選択肢も想定していることは、**サンプルセクションバイアスを考慮している**とも言える。前節ではパネルデータには一切欠測がない(**バランスドパネル**)ことを暗黙の前提としていたが、実際のパネルデータではしばしば欠測が起こる。もし欠測が企業の活動とは完全に無関係にランダムに発生している(MAR; missing at random)ならば推定に大きな影響はないが、すでに解説したように k_t は ω_t (の期待値)に相関し、 ω_t の大きさ次第で企業経営者は撤退を決める。つまりモデル内の変数に依存するメカニズムで発生している(MNAR; missing not →)ことになるので推定にバイアスを引き起こす^{*20}。

まず、推定のために投資が常に正である、 $inv_t > 0$ と仮定すると、投資関数 $inv_t(\cdot)$ が必ず単調増加関数になるので、

逆関数 $\omega_t = h_t(k_t, inv_t)$ が存在することになる。これを(2)に代入すると、

$$y_t = \beta_L l_t + [\beta_A + \beta_K k_t + h_t(k_t, inv_t)] + \varepsilon_t \quad (8)$$

となり、**観察されない効果を他の説明変数で表現できる**ことがわかった。 $[\dots]$ の部分は関数形が特定されていない、いわゆるノンパラメトリックモデルである。しかし、今知りたいのは $\beta_A, \beta_K, \beta_L$ であって、必ずしも投資関数を特定することは必須でない(このような推定方法を**コントロール関数アプローチ**という)。そこで、OP 法では以下のような**2段階の推定法**で一致推定量を得る。

第1段階では k_t と ω_t という動学的に決定される部分をノンパラメトリック回帰することで、静的に決まる β_L のみを識別する。

$$\phi_t(k_{i,t}, inv_{i,t}) := \beta_A + \beta_K k_{i,t} + h_t(k_{i,t}, inv_{i,t}) \quad (9)$$

と表せば、(8)は

$$y_t = \beta_L l_t + \phi_t(k_t, inv_t) + \varepsilon_t \quad (10)$$

と書き換えられる。 ϕ_t の部分をノンパラメトリック回帰で推定する^{*21}。これによって推定値 $\hat{\beta}_L, \hat{\phi}_t$ が識別された。これはいわゆる**部分線形モデル**である。

すると今度は ϕ_t からどうやって β_A, β_K を識別するかということを考えなければならない。

$$y_t - \hat{\beta}_L l_t = \phi_t(k_t, inv_t) + \eta_t$$

$\hat{\phi}_t$ には $\beta_A + \beta_K k_t + \omega_t$ が含まれているが、 k_t は既に説明したように動学的に決定される。ここで、次の $t+1$

^{*20} 欠測が推定結果に及ぼす影響については、星野(2009)、高井他(2016)、高橋・渡辺(2017)のいずれかが参考になるだろう。ただしパネルデータでの問題については星野が少し触れている程度である。英語でもいならばWooldridge(2010, Ch. 19)が解説している。

^{*21} Olley らは4次多項式補間で計算しているが、3次にしたりデータを結果に大きな変化がなかったとしている。一方で、年代で分割して試すと変化が見られたとしている。多項式に特定する理由もないので、他の方法も試すと良いだろう。

時点の $y_{t+1} - \beta_L l_{t+1}$ の期待値を考えると、以下ようになる (ω_t のマルコフ性を思い出す).

$$\begin{aligned} E[y_{t+1} - \beta_L l_{t+1} \mid k_{t+1}, x_{t+1} = 1] \\ = \beta_A + \beta_K k_{t+1} + g(\omega_{t+1}, \omega_t) \end{aligned}$$

ここで,

$$g(\omega_{t+1}, \omega_t) = \beta_A + E[\omega_{t+1} \mid k_{t+1}, x_{t+1} = 1]$$

と置いている. g は観測できない ω_t で構成されるが, (7) から, 退出確率は (ω_t, ω_{t+1}) の関数で書けることが分かっている. よって, g は p_t と ω_t で表現できる. $\phi_t = \beta_K k_t + \beta_A + \omega_t$ だから,

$$g(\cdots) = \beta_A + g(p_t, \phi_t - \beta_A - \beta_K k_t)$$

と書ける. 退出したかどうかは, パネルデータの欠測で知ることができるため, x_t に対してプロビット回帰をすることで退出確率の推定値 \hat{p}_t を計算できる. つまりこれは一種の**傾向スコア**である.

以上から, 第2段階^{*22}では, $\hat{\beta}_L, \hat{\phi}, \hat{p}_t$ を所与として, 以下のようなラグのある式を推定する. 非線形部分 $g(\cdots)$ はやはりノンパラメトリック回帰となる. これで β_A, β_K も推定できる.

$$(y_{t+1} - \hat{\beta}_L l_{t+1}) = \beta_K k_{t+1} + g(\hat{p}_t, \hat{\phi}_t - \beta_A - \beta_K k_t) + \xi_{t+1} + \eta_{t+1} \quad (11)$$

ここで, $\xi_{t+1} := \omega_{t+1} - E[\omega_{t+1} \mid \omega_t, x_{t+1} = 1]$ である. つまり ξ_{t+1} は平均ゼロで, かつ k_{t+1} と独立である. 一方で l_{t+1} とは独立していないが, β_L は第1段階で推定した $\hat{\beta}_L$ を代入できるので問題ない.

なお, OP 法では標準誤差を代数的に得られないため, ブートストラップシミュレーションで計算する必要がある.

4.5 OP 法以降の研究

Levinsohn and Petrin (2003) は OP の方法を改良したもの. 多くの企業の会計情報を見ると, 実際には設備投資は必ずしも発生せず, ゼロが多い. これは投資関数の単調性と矛盾するため, 中間投入財 (材料費) の情報を利用することを提案している. 現在はさらに Akerberg et al. (2015) が改良版を提案して, これがスタンダードになりつつあるらしいが今回は省略.

4.6 データのとり方

企業の生産関数推定で一番むずかしいのは実はここかもしれない. というのも, 理論上の K や L はどこにも記載されていないからだ. これはけっこう地道な計算が必要になる. 例えば企業の決算報告書などから地道に計算する, あるいは日銀や DBJ のような政策金融機関が整備しているデータベースをなんらかの手段で見せもらう, などである. あるいは, 農業や工業など資源と生産物の関係が分かりやすい特定の産業だけを分析の対象とすることもできる. 例えば農業で, K を乳牛の頭数, 原材料を飼料, Y をミルクの生産量とした分析も可能だろう (昔見た気がするがすぐにはみつけれなかった).

今回はとてもそんなことをしている暇も金もコネもない^{*23}ので, 後述する乱数データを使う.

^{*22} Olley らの論文では退出確率 p_t の推定を第2段階と呼んでいるので, これは本来は第3段階だった. しかし後の解説では省略されることが多い.

^{*23} いくつかの教科書の著者は練習問題用にデータセットを公開している. 例えば "Wooldridge data sets", "Datasets for empirical exercises in Econometrics" など. しかし, 今回の問題に適した企業別パネルデータは見つけられなかった.

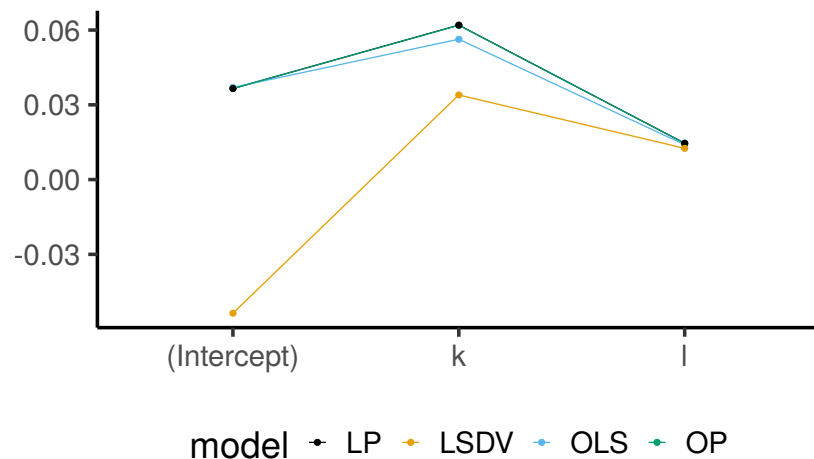


図2 推定結果

5 Rでの実装

実は Olley-Pakes や Levinsohn-Petrin の方法は R の [estprod](#) パッケージですでに実装されている。さらに、このパッケージには練習用の乱数データセットも用意されている。この乱数がどのような方法で生成されたかわからないので、実質的に動作確認にしかならないが、まずはこれを使ってみる^{*24}。時間があれば自分である程度実装したのだが、もう日付が変わっているので難しい。OLS と LSDV と比較した結果が図 2 である。LSDV が大きく異なる推定結果を出しているのがわかる。

5.1 モデルの事後診断

さて、反事実分析でのモデルの評価は単に当てはまりが良いかだけでなく、仮定と現実のデータが矛盾していないかの確認も同等かそれ以上に重要である。

ここでも Olley らの研究を例にする。まず、彼らの仮定が正しければ、単なる最小二乗法や、IV などの方法では結果にバイアスが発生するはずである。よって、提案する推定方法との比較が必要である。あるいは、サンプルセクションバイアスが実際に発生しているか確認するため、バランスドパネルでの結果と比較することも必要である。さらに、市場の構造は法規制の変化によって変わりうる。法改正のタイミングで年代ごとに区切って当てはまりを確認する必要がある。

さらに、モデルのロバストネス分析として、モデルの仮定のうち特に重要な、投資関数の式(6)を検討している。投資が資本と観察できない生産性によって決まるという仮定が正しくない場合、 β_L の推定にバイアスが生じる。これを検証するには最後の式(11)の説明変数に、現在の l_t を加えることのできる (このようなタイプの検証は**特定化のテスト**と呼ばれる)。

^{*24} このパッケージは CRAN に登録されているものの、作者のプロフィールがよくわからない。また、[github リポジトリ](#)にもほとんど関心が集まっていないため、動作がどの程度適切なのかあまり確認が進んでいなさそう。一応ソースコードを大まかに確認はしたが、その点は留意してほしい。あと、`coefs()`、`predict()`、といった stats の基本関数にも対応していない。もちろん `stargazer` に対応していないし、`estimatr` にも収録されていない。

6 まとめ

今回は取り掛かった段階で準備期間が1週間を切っていたため、かなり手を抜かざるを得なかった。Olley-Pakes や Levinsohn-Petrin は10年以上前の話なので、すでに大学院、あるいは学部の授業でも取り上げられることすらある。

ところで、香港科技大の教員であるKawaguchi (2019) が大学院生向けの講義ノートを一般公開している (公開期間限定かもしれない)。実はここでも、RでOP法を計算するプログラムを書けという課題が存在する。課題の数からして、これを受講する学生は毎週1回はこれくらいの課題をこなさなければならないようだ。これは経済学大学院ではよくある話 (伝聞によれば北米はもっと厳しい)。

参考文献

- Akerberg, Daniel A., Kevin Caves, and Garth Frazer (2015) "Identification Properties of Recent Production Function Estimators," *Econometrica*, Vol. 83, No. 6, pp. 2411–2451, DOI: [10.3982/ECTA13408](https://doi.org/10.3982/ECTA13408).
- Anderson, T. W. and Cheng Hsiao (1981) "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, Vol. 76, No. 375, pp. 598–606, September, DOI: [10.1080/01621459.1981.10477691](https://doi.org/10.1080/01621459.1981.10477691).
- (1982) "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, Vol. 18, No. 1, pp. 47–82, DOI: [10.1016/0304-4076\(82\)90095-1](https://doi.org/10.1016/0304-4076(82)90095-1).
- Arellano, Manuel and Stephen Bond (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies*, Vol. 58, No. 2, p. 277, April, DOI: [10.2307/2297968](https://doi.org/10.2307/2297968).
- Blundell, R and Steven Bond (1998) "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, Vol. 87, No. 1, pp. 115–143, November, DOI: [10.1016/S0304-4076\(98\)00009-8](https://doi.org/10.1016/S0304-4076(98)00009-8).
- Cameron, AC and PK Trivedi (2005) *Microeconometrics: Methods and Applications*, Cambridge: Cambridge University Press, DOI: [10.1017/CBO9781107415324.004](https://doi.org/10.1017/CBO9781107415324.004).
- Griliches, Zvi and Jacques Mairesse (1999) "Production Functions: The Search for Identification," in Strom, Steinar ed. *Econometrics and Economic Theory in the 20th Century*, Cambridge: Cambridge University Press, pp. 169–203, DOI: [10.1017/CCOL521633230.006](https://doi.org/10.1017/CCOL521633230.006).
- Hayashi, Fumio (2000) *Econometrics*, Princeton: Princeton University Press.
- Igami, Mitsuru (2018) "Artificial Intelligence as Structural Estimation: Economic Interpretations of Deep Blue, Bonanza, and AlphaGo," March, arXiv: [1710.10967](https://arxiv.org/abs/1710.10967).
- Kawaguchi, Kohei (2019) "ECON 6120I Topics in Empirical Industrial Organization," retrieved from [here](#).
- Levinsohn, James and Amil Petrin (2003) "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, Vol. 70, No. 2, pp. 317–341, April, DOI: [10.1111/1467-937X.00246](https://doi.org/10.1111/1467-937X.00246).
- Ljungqvist, Lars and Thomas J. Sargent (2018) *Recursive Macroeconomic Theory*, Cambridge, Mas-

- sachusetts: MIT Press, 4th edition.
- Lucas, Robert E. (1976) "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Confer. Series on Public Policy*, Vol. 1, No. C, pp. 19–46, DOI: [10.1016/S0167-2231\(76\)80003-6](#).
- Lütkepohl, Helmut (2007) *New Introduction to Multiple Time Series Analysis*, Berlin: Springer, 1st edition, OCLC: 254802195.
- Mairesse, Jacques (1990) "Time-Series and Cross-Sectional Estimates on Panel Data: Why Are They Different and Why Should They Be Equal?" in Hartog, Joop, Geert Ridder, and Jules Theeuwes eds. *Panel Data and Labor Market Studies*, North-Holland: Elsevier Science Publishers B.V., pp. 81–95.
- Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995) *Microeconomic Theory*: Oxford University Press.
- Merton, Robert C. (1974) "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *The Journal of Finance*, Vol. 29, No. 2, p. 449, May, DOI: [10.2307/2978814](#).
- Olley, G. Steven and Ariel Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, Vol. 64, No. 6, pp. 1263–1297, DOI: [10.2307/2171831](#).
- Reiss, Peter C. and Frank A. Wolak (2007) "Structural Econometric Modeling: Rationales and Examples from Industrial Organization," in Heckman, James J. and Edward E. Leamer eds. *Handbook of Econometrics*, Vol. 6 of Handbook of Econometrics: Elsevier, Chap. 64, pp. 4277–4415, DOI: [10.1016/S1573-4412\(07\)06064-3](#).
- Scott, Steven L. and Hal R. Varian (2014) "Predicting the Present with Bayesian Structural Time Series," *International Journal of Mathematical Modelling and Numerical Optimisation*, Vol. 5, No. 1/2, p. 4, DOI: [10.1504/IJMMNO.2014.059942](#).
- Stokey, Nancy L., Robert E. Lucas, and Edward C. Prescott (1989) *Recursive Methods in Economic Dynamics*, Cambridge, Mass: Harvard University Press.
- Varian, Hal R. (1992) *Microeconomic Analysis*, New York: Norton, 3rd edition, (佐藤隆三・三野和雄訳, 『ミクロ経済分析』, 勁草書房, 1986年) .
- Wooldridge, Jeffrey M (2010) *Econometric Analysis of Cross Section and Panel Data*: The MIT Press, 2nd edition, retrieved from [here](#).
- 阿部修人 (2017) 「上級マクロ経済学講義ノート動的計画法」, retrieved from [here](#).
- 沖本竜義 (2010) 『経済・ファイナンスデータの計量時系列分析』, 朝倉書店.
- 奥井亮 (2016) 「動学的パネルデータモデル」, retrieved from [here](#).
- 北川源四郎 (2005) 『時系列解析入門』, 岩波書店.
- 北村行伸 (2016) 「応用ミクロ計量経済学的手法と論点」, 『経済セミナー増刊進化する経済学の実証分析』, 日本評論社, 28–36 頁, retrieved from [here](#), 個人サイトにリンクないんだけど落として良かったのか?.
- 北村行伸・西脇雅人・村尾徹士 (2009) 「不完全資本市場下での生産関数の推定について」, 一橋大学ディスカッションペーパー 70, 193–204 頁, retrieved from [here](#).
- 工藤教孝 (2007) 「動学的最適化入門」, Technical report, retrieved from [here](#).
- 久保拓哉 (2012) 『データ解析のための統計モデリング入門』, 岩波書店.
- 高井啓二・星野崇宏・野間久史 (2016) 『欠測データの統計科学』, 岩波書店.
- 千木良弘朗・早川和彦・山本拓 (2011) 『動学的パネルデータ分析』, 知泉書館, 東京, OCLC: 751998792.
- 中島上智・渡部敏明 (2012) 「時変ベクトル自己回帰モデル-サーベイと日本のマクロデータへの応用-」, 『経済

- 研究』, 第 63 卷, 第 3 号, 193–208 頁, DOI: [10.15057/25864](https://doi.org/10.15057/25864).
- 中嶋亮 (2016) 「誘導型推定」 v.s. 「構造推定」, 『経済セミナー増刊進化する経済学の実証分析』, 日本評論社, 52–62 頁.
- 西村和雄 (1990) 『ミクロ経済学』, 東洋経済新報社.
- 蓮見亮 (2019) 「動学マクロ経済学入門」, retrieved from [here](#).
- 星野崇宏 (2009) 『調査観察データの統計科学 – 因果推論・選択バイアス・データ融合』, 岩波書店.
- 渡部敏明 (2016) 「ルーカス批判とマクロ計量分析」, 『経済セミナー増刊進化する経済学の実証分析』, 日本評論社, 37–41 頁.
- 高橋将宜・渡辺美智子 (2017) 『欠測データ処理 – R による単一代入法と多重代入法–』, 統計学 One Point, 第 5 号, 共立出版.