

経済学と反事実分析 接触篇

Economics and Counterfactual Analysis: A Contact

ill-identified

Tokyo.R 83rd, 2020/1/25, Updated: 2020/1/26

twitter: **ill-identified**



- (~2014): 学生 (経済学)
- (~2017): だいたい SAS エンジニア
- (2018~): アドテク, 機械学習エンジニア
- 詳しい経歴: **LinkedIn**

自己紹介

- 主語の**大きい**話をするのが好きです
 - 『計量経済学と機械学習の関係』 (80th Tokyo.R)
 - 『今更だが, ベイズ統計とは何なのか.』
- 主語の**変な**話をすることもあります
 - 『三国志で学ぶデータ分析』 (Japan.R '19)
 - 『飯野山は正規分布らしいのでパラメータを推定する』

【PR】勤務先の求人

FANCOM

TOP

BLOG

RECRUITMENT

データサイエンティスト

インターネット広告配信システムのデータ分析、配信ロジックの考案、効果測定と改善まですべての工程に携わっていただきます。広告配信で蓄積された膨大なデータを活用し、ユーザーにとってはより最適な広告を、広告主にとってはより効果的な広告を配信できるようにすることがミッションです。

内容

- インターネット広告配信システムのデータ分析
- 配信ロジックの考案
- 広告効果の測定と改善

スキル

下記のいずれか

- プログラミングでの実装経験（言語問わず）
- 統計学の知識、経験
- 画像認識技術の知識、経験
- ディープラーニングの知見
- 機械学習やデータサイエンス関連知識、経験
- SQL、Python、R、SAS、SPSSの実務経験
- 関連論文の情報収集、実装経験

AGENDA

- 前は因果推論メインなので構造推定やる
- 結構昔の話なので知ってる人もいるかも

介入実験と反事実実験

生産関数の推定

構造推定の具体例

R でやる

まとめ

発表後の反響に対する補足説明

- スクリプトと原稿はここ

介入実験と反事実実験

構造推定とは 1

- 構造方程式モデリング (SEM, aka 共分散分析)
 - ダイアグラムを特定して $a_{ij} \neq 0$ を求める

$$X_1 = a_{1,1}X_1 + a_{1,2}X_2 + \cdots + a_{1,p}X_p + e_1$$

$$X_2 = a_{2,1}X_1 + a_{2,2}X_2 + \cdots + a_{2,p}X_p + e_2$$

\vdots

$$X_p = a_{p,1}X_1 + a_{p,2}X_2 + \cdots + a_{p,p}X_p + e_p$$

- ... ではない

構造推定とは 2

- 構造ベクトル自己回帰モデル (Structural VAR)
 - 同時決定を考慮した多変量時系列モデル

$$\mathbf{A}\mathbf{y}_t = \mathbf{B}_0 + \mathbf{B}_1\mathbf{y}_{t-1} + \cdots + \mathbf{B}_p\mathbf{y}_{t-p} + \mathbf{e}_t$$

- ... でもない

構造推定とは 3

- 第 32 回 Tokyo.R (2012 年) 「構造型モデル」
- R で学ぶ『構造型モデル de 倒産確率推定』
- ... **これも違う**
 - 経済学的な観点では構造を考えている
 - 細かい話をするとうの長さになるので省略

構造推定とは 4

1. 経済理論に基づいてモデルを作る
2. 観察データからモデルの**構造パラメータ**を推定
3. モデルから**因果関係**を考えることができる
 - [19, 26, 25, 22] あたりが分かりやすい解説?
 - テクニカルな話:[29, 10, 23]

お前は構造推定じゃない

- 構造方程式: 経済理論に基づいてない
 - 参考: [28]
- 構造 VAR: 経済学だけど経済モデル明示してない
 - 参考: 沖本 [20] (沼: [12])
- 構造型モデル: 最適化行動だが均衡概念を考慮してない
 - 参考: Merton モデル [14]

反事実と介入

- 因果推論 (RCT, IV, DID, RDD...) との違いは?
 - 因果推論: 観察したデータのみに基づく (擬似的な反事実)
 - 構造推定: データモデルから**現実と異なる状況**を分析可能 (思考実験的?)
- 因果推論 = **介入** (intervention) または**プログラム評価**, 構造推定 = **反事実** (counterfactual)

構造形と誘導形

- 経済学では**構造形**・**誘導形**という言い方も
 - 構造 VAR と誘導形 (SUR)

$$\mathbf{A}\mathbf{y}_t = \mathbf{B}_0 + \mathbf{B}_1\mathbf{y}_{t-1} + \cdots + \mathbf{B}_p\mathbf{y}_{t-p} + \mathbf{e}_t$$

$$\mathbf{y}_t = \mathbf{A}^{-1}\mathbf{B}_0 + \mathbf{A}^{-1}\mathbf{B}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}^{-1}\mathbf{e}_t$$

- SUR 推定では $\mathbf{A}, \{\mathbf{B}_i\}$, が不明 (予測結果は同じ)
- 因果推論も構造形を変形した誘導形
 - 例: 操作変数法, 2 段階最小二乗法

どっちが強いのか

- 2010 年前後に学会で**激論** [25]
 - 誘導形派「因果推論はエビデンスベースド, 構造推定は恣意的な仮定」
 - 構造形派「因果推論は仮定をごまかしているだけ, 結果を一般化できない」
- 現在は各々できることの限界がより明確に
- 構造推定は**事前評価**, 因果推論は**事後評価** (山口 [29])

いろいろな構造推定

- 経済モデル次第でやり方は不定形
 - 経済学の数学テクニックを総動員する
- 以下は便宜的な分類
- **動学的**構造推定
 - 伊神 [9]「**構造推定 = AI**」(勝手な解説1, 2)
 - 時間的な変化を考慮するため計算が多く大変
- **静学的**構造推定
 - **需要関数の推定** [4], **生産関数の推定** [15], [11]
 - 時間変化を意識せず計算できることが多い

生産関数の推定

生産関数の推定

- コブ = ダグラス型生産関数

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

- 資本 (K) と労働力 (L) と生産量 (Y) の関係

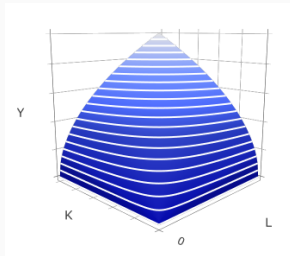


図 1: コブ = ダグラス型関数の例

どう推定するか

- 対数を取ればただの重回帰:

$$\ln Y_i = \ln \alpha + \beta_K \ln K_i + \beta_L \ln L_i + u_i$$
$$\Leftrightarrow y_i = \beta_A + \beta_K k_i + \beta_L l_i + \varepsilon_i$$

- ... **ではない**

経済学・統計学的に見てだめな理由

1. 企業特有の効果を観察不可能
 - ・ 結果にバイアス
2. 需要要因を考慮していない
 - ・ もし需要が変化しても同じこと言えんの？
3. セレクションバイアス
 - ・ 企業は立ち上がったり倒産したり

企業特有の効果とパネルデータ

- 企業固有の生産性 (企業文化, 謎の超技術 etc.) ω_i はデータとして**観測不可**

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

- 各企業を追跡調査した**パネルデータ**を使う

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

パネルデータ分析の教科書

- (昔アイコンに使っていた赤いやつは初版)

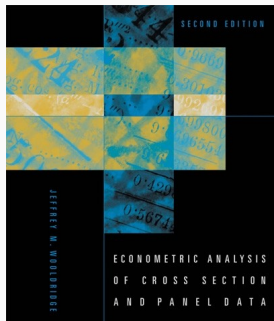


図 2: Wooldridge の教科書 [17]

パネルデータの個別効果モデル

- 企業ごとにダミー変数 (LSDV)

$$y_{i,t} = \beta_A + \beta_K k_{i,t} + \beta_L l_{i,t} + d(1)_i + \cdots + d(N)_i + \varepsilon_{i,t}$$

- 企業毎平均を引く (Fixed Effect, Within)

$$y_{i,t} - \bar{y}_i = \beta_A + \beta_K (k_{i,t} - \bar{k}_i) + \beta_L (l_{i,t} - \bar{l}_i) + \cdots$$

- Wooldridge Ch.10[17], 奥井の講義スライド[21] が参考に

補足: GLMM と個別効果

- 久保 [24] の一般化線形混合モデル (GLMM)
 - 個別効果モデルの一般形
- 線形モデルは平均除去して推定できるという話



図 3: 久保本の表紙

個別効果の限界

- 個別効果 ω は他の変数と**無相関**という前提
(**内生性**の問題)
- 生産関数の推定には不十分と指摘 [13]
- **動学パネルデータモデル**

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

- Anderson-Hsiao[1, 2], GMM 推定 [3, 5]
 - より過去の $y_{i,t-2}, \dots$ なら無相関なので操作変数に
- ... **雑すぎない?**

構造推定の実例

生産関数の構造推定

- ω と説明変数の相関は企業経営者の意思決定メカニズム (内生性)
- Olley と Pakes[15] の構造推定を紹介
- ... **ここからが本題**

モデルの仮定

- 一旦 t の変化だけ考える

$$y_t = \beta_A + \beta_K k_t + \beta_L l_t + \omega + \varepsilon_t$$

- ω_t は企業固有の効果だが, 外部要因の影響も (一般均衡)

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

- 経営者は ω_t を見て今期経営計画を決定
 - 設備投資 (inv_t) をどうするか (k_t に影響)
 - 事業を撤退すべきか**
 - 労働力 (l_t) は独立

わかりにくいので図で

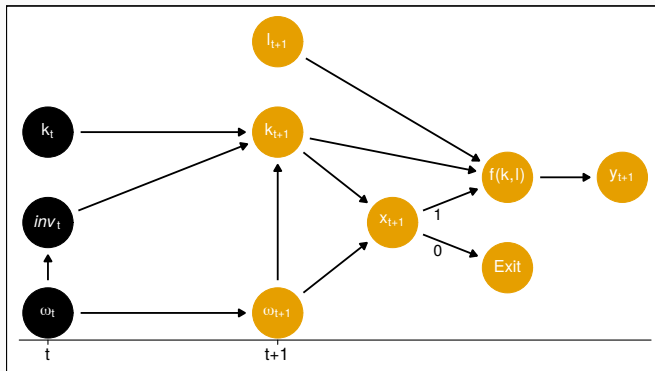


図 4: OP モデルで仮定した変数の関係

動学最適化

- 経営者は将来を見越して再帰的に意思決定
- OP 法は最適化問題 (ベルマン方程式) を計算せずにモデルを推定

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

- この式は覚えなくてもいい

推定方法

1. $\beta_K k_t + \omega_t$ が相関するなら ϕ_t で 1 まとめにして $\hat{\beta}_L$ だけ求める

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

- ϕ_t には多項式回帰, カーネル回帰などを使う

2. 撤退確率 p_t をプロビットで推定
3. ϕ_t から ω_t を取り除く. $\hat{\phi}_t, \hat{p}_t, \hat{\beta}_L$ を使った $\hat{\omega}_t$ で置換 (t のズレに注意)

$$(y_{t+1} - \hat{\beta}_L l_{t+1}) = \beta_K k_{t+1} + g(\hat{p}_t, \hat{\omega}_t) + e_{t+1}$$

- $g(\cdot)$ もノンパラメトリック回帰

統計モデル的な意味

- k_t の内生性に対処
 - コントロール関数アプローチ
- p_t の計算
 - サンプルセレクションバイアスを傾向スコアで補正
 - パネルデータの欠測を補正

R でやる

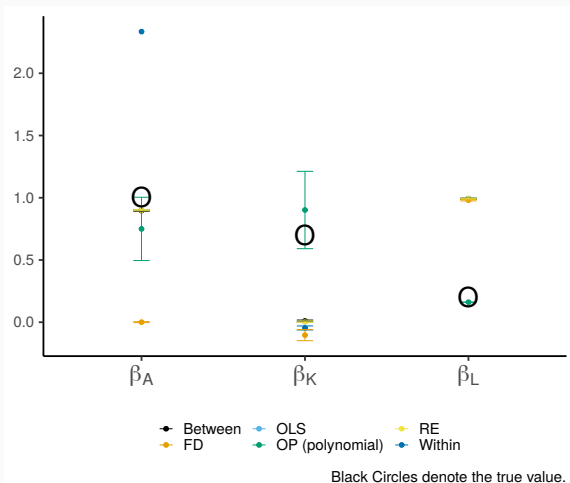
R でやる

- パネルデータは **AER**, **plm** がある (参考)
 - 最近使っていないのでよくわからん
- ... **自作する時間がない**
- **estprod** パッケージがある
 - 現状使い勝手は良くない
 - データ変えて検証したらうまく行かなかった
ので自作する

```
olley_pakes(y ~ l | k | inv, data = df,  
  exit = ~exit, id = "id", time = "year",  
  bootstrap = T)
```

推定結果

- Arellano-Bond 等ズレの大きいモデル省略



まとめ

構造推定を計算した後...

- 構造推定は計算して終わりではない
- 仮定の正しさ確認 (ロバストネス分析)
 - データを分割して推定
 - **仮説と矛盾するモデル**もあえて推定する (特定化テスト)
- 上記をパスしたら反事実的な主張をする
 - OP 論文はパラメータ推定がメインでそこまで大胆な主張してない
 - **ぶっちゃけ「反事実」の典型例とは言えない**

ところで...

- 香港科技大の博士課程向け授業の講義ノート
 - (川口先生作)[10]
 - 課題 2 に LP 法をやれというものがある
 - 課題の数的に猶予は 1 週間くらい?
 - 挑戦してみてください

まとめ

- 構造推定は経済理論モデルとデータの融合
- 構造推定は反事实的分析ができる
- 因果推論と構造推定に各々の向き不向き
- **estprod** で OP 法の基本部分は計算できる
- ... **じかんがない**
- 次回から簡単で実用的な話を増やす
- (追加) OP 法は技術的には簡単な部類だが
目的意識がわかりにくかった

経済学と反事実分析

接触遅刻篇

完

発表後の反響に対する補足説明

復習：線形回帰モデルの基本

1. $\beta_A, \beta_K, \beta_L$ の不偏推定の条件 (外生性の仮定)

$$E[\omega_t + \varepsilon_t \mid k_t, l_t] = 0$$

2. (1) の必要条件 (直交条件)

$$E \begin{bmatrix} k_t & l_t \end{bmatrix} (\omega_t + \varepsilon_t) = 0$$

- パネルデータは直交条件前提
 - ω_t が**平均ゼロ**かどうかが焦点
- 今回は ω_t が k_t に相関
 - 明らかに**直交条件も満たさない**
- 正確な説明:[18], [8], [17]

まじめに実装: 乱数データの作成

- estprod の練習データは生成方法が謎
- 講義ノート [10] を参考に乱数データ生成
- 退出確率がオミットされているので追加

1 段階目の推定

- np パッケージで部分線形モデルを推定可能
- npplregbw (バンド幅探索), npplreg (回帰)
 - **注:** なぜか `na.action` が効かない
- 多項式近似は `lm()` と `poly()`
- プロビット回帰
 - ロジスティック回帰を正規分布に置き換え
 - `glm(family = binomial(link = "probit"))`

2 段階目の推定

- t のラグ項が必要
 - `dplyr::lag()` でデータフレームを加工し対処
- `nls()` 関数で**非線形最小二乗法**による計算
- 課題 [10] は簡略化のため**非線形 GMM**で推定可
 - `gmm::gmm()` 関数で推定.
 - GMM の話は [18], [8], 前書いたやつ
 - `plm` の GMM(Arellano-Bond, Blundel-Bond) は**非線形不可**.

補足: nls 関数の使い方

- formula でモデル記述
- シンボル名は入力データかパラメータかのどちらか
- データにない = パラメータ, 初期値必須
- 例: データ (x_i, y_i) に対して $y_i = (x_i - a)^2$ の a を求めたい:
- 多項式展開の記述が大変, optim() も可

```
nls(y ~ (x-a)^2, data, start = list(a = 1))
```

補足: ブートストラップ法

- モンテカルロ法めいている
 - 詳細は [6, 7]
- 元データからランダムサンプリングで擬似的に複数のデータセット作成
- 各データの推定結果の平均や分散を計算
 - つまり標準誤差を計算できる
- 回帰モデルの場合は少し複雑
 - 予測値 \hat{y}_i と残差 $\hat{\epsilon}_i$ に分解し新しいデータ $\tilde{y}_i = \hat{y}_j + \hat{\epsilon}_k$ を作成

補足: boot パッケージ

- ブートストラップ法を計算するパッケージの 1 つ
- `boot()` 関数にデータ, 欲しい統計量 (推定量) を計算する関数, を与える
- 関数はランダムサンプリングを表すインデックスを受け取れるようにする
- 細かいオプションに関する解説: 『**ブートストラップのための boot パッケージ**』.

補足: boot パッケージの用例

- 線形回帰の係数をブートストラップ法で

```
fit <- lm(y ~. data = df)
df <- mutate(df, fitted = fitted(fit), resid
             = y - fitted)
stat <- function(data, index){
  data$y <- data$fitted + data$resid[index]
  fit <- lm(y ~. data = data)
  return(coef(fit))}
bs <- boot(df, stat, R = 100)
```


補足: gmm 関数の使い方

- 線形モデル: formula で記述
- 非線形モデル: 標本モーメント
 $(y_i - f(x_i)) \times [z_1 \cdots z_q]$ の行列を返す関数
を与える
- 目的関数を自作すれば optim() で同じ係数
を再現できるはず

再確認: Olley と Pakes の問題意識

- 目的の 1 つは**全要素生産性** (TFP) 分析
 - 法改正が生産性にどう影響するか
 - TFP は回帰分析の残差で計測 (**ソロー残差**)
- 個別効果の推定方法を複数試す [13]
 - 全部係数がばらばら
 - 生産メカニズムを特定できていない疑惑
 - その原因は**何らかの内生性**

再確認: 変数のカテゴリ

- 形式的に変数は 3 つのグループ

1. 外生変数: ω_t と相関しない変数 (e.g., l_t),
2. 内生変数: ω_t に依存する変数 (e.g., k_t),
3. 操作変数: ω_t と上記を識別する変数 (e.g., inv_t)

当てはまりの良さ \neq 分析の正しさ

- 生産関数の本当の構造は誰にも分からない
 - TFP= 理論モデルで説明できない部分
- 当てはまりではなく係数の一致推定が重要
 - 「当てはまりが良い**だけ**」は問いへの答えではない
- 分析は必ず「ある仮定のもとで」という前置きが存在
 - cf. カール = ポパーの「反証可能性」
 - cf. 「構造形」vs「誘導形」論争

その他の構造推定研究事例

- 鉄道の混雑に対する人々の行動分析 [27]
 - 人間は混雑を予測して行動している
 - 観察データから推定しただけなら結果にバイアス
- …これは (動学) 構造推定…!!
 - ゲーム理論でモデル化, データに基づき推定
 - 計算方法も過去の構造推定論文の応用
- 動学構造推定は計算が大変なのでアルゴリズムの理解も必要

構造推定は「役に立つ」のか？

- 構造モデルを大規模な実験で検証 [16]
- 「ネット広告オークションに**留保価格**は必要？」
 - 留保価格決定方法を考案
 - オークション理論・ゲーム理論の構造推定
- 米 Yahoo! 「業績の向上はこの研究によるところが大きい」
 - 単純に RCT(事後評価) するだけでは不可能
- もちろんどう役に立たせるかは人間しだい

参考文献

- [1] 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).
- [2] ——— (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).
- [3] 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).
- [4] Berry, Steven, James Levinsohn, and Ariel Pakes (1995) "Automobile Prices in Market Equilibrium," *Econometrica*, Vol. 63, No. 4, p. 841, July, DOI: [10.2307/2171802](https://doi.org/10.2307/2171802).

- [5] 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).
- [6] Efron, B. (1979) "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics*, Vol. 7, No. 1, pp. 1–26, January, DOI: [10.1214/aos/1176344552](https://doi.org/10.1214/aos/1176344552).
- [7] Efron, Bradley and Robert Tibshirani (1993) *An Introduction to the Bootstrap* in , Monographs on Statistics and Applied Probability, No. 57, New York: Chapman & Hall.
- [8] Hayashi, Fumio (2000) *Econometrics*, Princeton: Princeton University Press.
- [9] 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).

- [10] Kawaguchi, Kohei (2019) "ECON 6120I Topics in Empirical Industrial Organization," retrieved from [here](#).
- [11] 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](#).
- [12] Lütkepohl, Helmut (2007) *New Introduction to Multiple Time Series Analysis*, Berlin: Springer, 1st edition, OCLC: 254802195.
- [13] 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.

- [14] 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](https://doi.org/10.2307/2978814).
- [15] 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](https://doi.org/10.2307/2171831).
- [16] Ostrovsky, Michael and Michael Schwarz (2016) “Reserve Prices in Internet Advertising Auctions: A Field Experiment,” Working Paper 2054, Stanford Graduate School of Business, Stanford, CA, USA, p. 23, retrieved from [here](#).
- [17] Wooldridge, Jeffrey M (2010) *Econometric Analysis of Cross Section and Panel Data*: The MIT Press, 2nd edition, retrieved from [here](#).
- [18] 浅野哲・中村二郎 (2009) 『計量経済学』, 有斐閣, 第 2 版.

- [19] 今井晋・有村俊秀・片山東 (2001) 「労働政策の評価: 「構造推定アプローチ」と「実験的アプローチ」」, 『日本労働研究雑誌』, 第 497 号, retrieved from [here](#).
- [20] 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.
- [21] 奥井亮 (2016) 「動学的パネルデータモデル」, retrieved from [here](#).
- [22] 北村行伸 (2016) 「応用ミクロ計量経済学の手法と論点」, 『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?.
- [23] 楠田康之 (2019) 『経済分析のための構造推定アルゴリズム』, 三恵社, 名古屋, retrieved from [here](#), OCLC: 1129899352.
- [24] 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.

- [25] 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.
- [26] 楡井誠 (2010) 「構造推定・一般均衡と政策評価」, retrieved from [here](#).
- [27] 松村杏子・武藤滋夫・福田大輔・柳沼秀樹 (2008) 「混雑した都市鉄道における出発時刻選択モデルの構造推定」, 『土木計画学研究・講演集』, 第 45 巻, 第 165 号, retrieved from [here](#).
- [28] 宮川雅巳 (2004) 『統計的因果推論–回帰分析の新しい枠組み–』, シリーズ予測と発見の科学, 第 1 号, 朝倉書店, Tokyo, retrieved from [here](#).
- [29] 山口慎太郎 (2017) 「動学的離散選択モデルの構造推定」, 『第 20 回労働経済学カンファレンス』, 東京, 9 月, retrieved from [here](#).