

# 予測研究: まとめ

経済学のための機械学習入門

川田恵介

## Table of contents

応用事例	1
同時手番 2 Player ゲーム . . . . .	2
ゲーム理論 VS 教室つき学習 . . . . .	2
結果 . . . . .	2
総まとめ	2
Well-specified Linear Model . . . . .	3
予測誤差の分解 . . . . .	3
教科書的な性質 . . . . .	3
Bias-Variance decomposition . . . . .	3
Miss-specified Linear Model . . . . .	4
Best Linear Projection . . . . .	4
母集団上での記述統計量の推定 . . . . .	4
Miss-specified Linear Model . . . . .	4
More complicated Model . . . . .	5
伝統的なトレードオフ . . . . .	5
Pruned Tree/Random Forest . . . . .	5
まとめ . . . . .	6
Reference . . . . .	6

## 応用事例

- “経済モデル” を予測性能の観点から評価
  - 伝統的なアプローチだが、しばしば”Naive” にやられてきた
- (Fudenberg and Liang 2019) : ゲーム理論 (同時手番 2 Player) への応用

- Sahoo, Lei, and Wager (2022) , Andrews et al. (2022)

## 同時手番 2 Player ゲーム

- 二人のプレイヤーが同時に” 戦略  $\{1, 2, 3\}$  “を選ぶ
- 戦略の組み合わせによって、利得を得る
  - 利得票

	b1	b2	b3
a1	"3,3"	"4,1"	"1,3"
a2	"3,3"	"6,6"	"1,1"
a3	"6,3"	"3,5"	"3,2"

- 利得 ( $= X$ ) から、選択される戦略 ( $= Y$ ) を予測できるか?

## ゲーム理論 VS 教室つき学習

- ゲーム理論は、さまざまな均衡予測を提供
  - 予測 + その理由 (解釈) を提供
  - 同じゲームを繰り返すと、理論予測に収束する場合も多い
  - 最初にプレイされるゲームの戦略は予測が難しい
- 純粋に予測のみを目的とした教師付き学習の予測にどこまで (どうすれば) 迫れるか?
  - そもそも予測が不可能な状況を” 排除” できる

## 結果

- 実験室実験のデータについて予測を行うと
  - 伝統的な均衡概念 (Uniform Nash, Level-1) は、Bagging による予測に負ける
  - Level-1 は、Nash 均衡に比べて、かなり Bagging に迫る予測を得られる

## 総まとめ

- 計量経済学入門: 「Well-specified Model を推定」が前提
- Miss-specified Model の可能性を考慮すると?
  - 経済学の応用において極めて重要な論点 (Gelman and Vehtari 2021)

- 教師付き学習: 十分に複雑なモデルからスタートすることで、Miss-specification の可能性緩和

## Well-specified Linear Model

- $\beta$  を適切に選べば、 $E_P[Y|X] = g(X)$
- 一致性、普遍性を満たす
  - 中心極限定理より、ある程度の事例数があれば、推定値は正規分布で近似できる (漸近正規性)
- 手近な学部中級以上のテキストで確認可能

## 予測誤差の分解

$$\begin{aligned}
 Y - g(X) &= \underbrace{Y - E_P[Y|X]}_{\text{Irreducible}} \\
 &+ \underbrace{E_P[Y|X] - g_\infty(X)}_{\text{Approximation Error}} \\
 &+ \underbrace{g_\infty(X) - g(X)}_{\text{Estimation Error}}
 \end{aligned}$$

## 教科書的な性質

$$\begin{aligned}
 Y - g(X) &= Y - E_P[Y|X] \\
 &+ \underbrace{E_P[Y|X] - g_\infty(X)}_{=0 : \text{Consistency}} \\
 + \underbrace{g_\infty(X) - g(X)}_{\sim N(0, \sigma^2) : \text{Asymptotic Normality}}
 \end{aligned}$$

## Bias-Variance decomposition

$$\begin{aligned}
 Y - g(X) &= Y - E_P[Y|X] \\
 &+ \underbrace{E_P[Y|X] - g_\infty(X)}_{=0 \text{ Consistency}}
 \end{aligned}$$

$$+ \underbrace{g_{\infty}(X) - E[g(X)]}_{=0 : Bias}$$

$$\underbrace{E[g(X)] - g(X)}_{\sim N(0, \sigma^2) : Variance}$$

## Miss-specified Linear Model

- $\beta$  をどう選んでも、 $E_P[Y|X] \neq g(X)$
- $g_{\infty}(X) = ?$  (何を推定しているのか?)
  - 母集団上で定義できる (研究者間で合意可能な) 推定ゴールは存在するのか?

## Best Linear Projection

$$g_{\infty}(X) = g^{BLP}(X) := \beta_0^{BLP} + \dots + \beta_L^{BLP} X_L$$

- ただし

$$\min_{\beta_0^{BLP}, \dots, \beta_L^{BLP}} E_P[(Y - g^{BLP}(X))^2]$$

- 母集団上での仮想的な回帰結果
  - 上級レベルの計量経済学テキスト (Hayashi や Wooldreige など) では紹介

## 母集団上での記述統計量の推定

- OLS = 母集団の記述統計である BLP の推定と解釈
  - 研究者が設定したモデルに依存するが、母集団上で定義可能
  - 母分布  $f_P(Y, X)$  の部分的に要約する
  - 母平均  $E_P[Y]$  : BLP の特殊ケース  $g_{BLP}(X) = \beta_0^{BLP}$
- 予測研究への応用: 母集団の記述統計量を推定する手法を流用

## Miss-specified Linear Model

$$Y - g(X) = Y - E_P[Y|X]$$

$$\begin{aligned}
& + \underbrace{E_P[Y|X] - g_{BLP}(X) (= g_\infty(X))}_{\neq 0} \\
& + \underbrace{g_{BLP}(X) - E[g(X)]}_{=0} \\
& + \underbrace{E[g(X)] - g(X)}_{\sim N(0, \sigma^2)}
\end{aligned}$$

More complicated Model

$$\begin{aligned}
Y - g(X) &= Y - E_P[Y|X] \\
& + \underbrace{E_P[Y|X] - g_{BLP}(X)}_{\neq 0} \\
& + \underbrace{g_{BLP}(X) - E[g(X)]}_{=0} \\
& + \underbrace{E[g(X)] - g(X)}_{\text{増加}}
\end{aligned}$$

伝統的なトレードオフ

- Approximation Error VS Variance
- 教師付き学習の多くの手法とは異なる
  - 緩やかな条件で一致性が成り立つ (RandomForest/決定木については、Chi et al. (2022) とその引用文献などを参照)
  - Bias VS Variance
- 十分に複雑なモデルから始めると、Approximation Error  $\sim 0$

Pruned Tree/Random Forest

$$\begin{aligned}
Y - g(X) &= Y - E_P[Y|X] \\
& + \underbrace{E_P[Y|X] - g_\infty(X)}_{\sim 0}
\end{aligned}$$

$$+ \underbrace{g_{\infty}(X) - E[g(X)]}_{\neq 0}$$

$$+ \underbrace{E[g(X)] - g(X)}_{\text{減少}}$$

## まとめ

- OLS 推定の解釈は、母集団の”記述統計量 BLP”を推定していると一般化可能
  - 「Well-specification である」という特殊ケースにおいて、条件付き平均値を推定している
- 最悪でも”BLP”という母集団で定義可能かつ“解釈可能”な値を推定しているので、今でもよく使われる
- 最尤法についてもよく似た解釈は可能だが、ベイズは？ (Buja, Brown, Berk, et al. 2019; Buja, Brown, Kuchibhotla, et al. 2019)

## Reference

- Andrews, Isaiah, Drew Fudenberg, Annie Liang, and Chaofeng Wu. 2022. “The Transfer Performance of Economic Models.” *SSRN Electronic Journal*.
- Buja, Andreas, Lawrence Brown, Richard Berk, Edward George, Emil Pitkin, Mikhail Traskin, Kai Zhang, and Linda Zhao. 2019. “Models as Approximations i.” *Statistical Science* 34 (4): 523–44.
- Buja, Andreas, Lawrence Brown, Arun Kumar Kuchibhotla, Richard Berk, Edward George, and Linda Zhao. 2019. “Models as Approximations II.” *Statistical Science* 34 (4): 545–65.
- Chi, Chien-Ming, Patrick Vossler, Yingying Fan, and Jinchi Lv. 2022. “Asymptotic Properties of High-Dimensional Random Forests.” *The Annals of Statistics* 50 (6): 3415–38.
- Fudenberg, Drew, and Annie Liang. 2019. “Predicting and Understanding Initial Play.” *American Economic Review* 109 (12): 4112–41.
- Gelman, Andrew, and Aki Vehtari. 2021. “What Are the Most Important Statistical Ideas of the Past 50 Years?” *Journal of the American Statistical Association* 116 (536): 2087–97.
- Sahoo, Roshni, Lihua Lei, and Stefan Wager. 2022. “Learning from a Biased Sample.” *ArXiv* abs/2209.01754.