

Sensitivity Analysis

川田恵介

2025-07-27

1 不適切な推論

1.1 観察されない変数を求める研究課題

- 労働経済学の研究目標はしばしば、観察できない変数も含めた母分布の推定が求められる
 - ▶ 因果効果: 技能訓練の効果を知りたい
 - 訓練への参加がランダムに決まるためには、“キャリアの志向”もバランスさせたいが分からない
 - ▶ 格差: 同じ教育経験を有する労働者内で、男女間賃金格差を知りたい
 - 多くのデータは、教育年数などの、“荒い”変数しか分からない

1.2 推定目標

- 本来であれば、データから観察できる変数 X に加えて、観察できない変数 U (キャリア志向、細かい教育経験) もバランスしたい
 - ▶ $Y \sim D + X + U$
を OLS 推定し、 β_D を推定したい
- 実際には、 $Y \sim D + X$ しか推定できず、 β_D^{short} しか推定できない

1.3 Sensitivity 分析

- $Y \sim D + X$ の推定結果から、 $Y \sim D + X + U$ の結果を類推できないか?
 - ▶ U の影響について、シナリオを設定する

1.4 Cinelli & Hazlett (2020)

- 入門 + R/STATA での実装法紹介 (Cinelli et al., 2024)
- U の影響を 2 つの指標で捉える
 - ▶ Y についての追加的な説明力 (Partial R^2): $R^2_{Y \sim U|D,X}$

- D についての追加的な説明力: $R_{D \sim U|X}^2$

1.5 推定値の関係性

- 推定値 $\hat{\beta}_D, \hat{\beta}_D^{short}$ の間に、以下の関係性が成り立つ

$$|\hat{\beta}_D - \hat{\beta}_D^{short}| = \sqrt{\frac{R_{Y \sim U|D,X}^2 \times R_{D \sim U|X}^2}{1 - R_{D \sim U|X}^2}} \times constant$$

- constant は、 $\hat{\beta}_D^{short}$ の標準誤差と事例数(自由度)に依存
 - U の Y, D について説明力が大きければ、乖離が大きい

1.6 例: 学歴の効果

```
data("CPSSW8", package = "AER")

model <- lm(earnings ~ education + gender + age, CPSSW8)

model
```

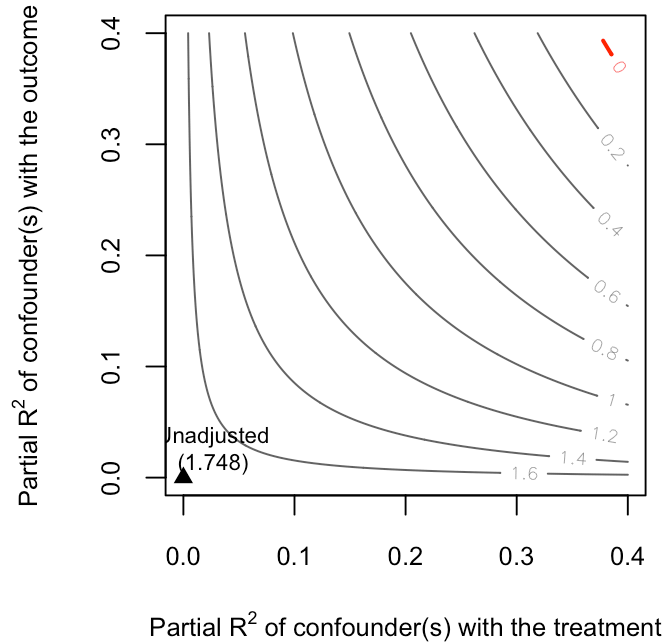
```
Call:
lm(formula = earnings ~ education + gender + age, data = CPSSW8)

Coefficients:
(Intercept)      education  genderfemale          age
    -10.0257         1.7481        -4.2502         0.1572
```

- U = 親の社会/経済状況

1.7 例

```
sensemakr::ovb_contour_plot(
  model,
  treatment = "education"
)
```



1.8 ベンチマーク

- $R_{Y \sim U|D,X}^2$, $R_{D \sim U|X}^2$ の大きさについて、どのような想定をすれば良いか?
- 一案は、 X の中から、 U との関係性が強そうな変数を回帰から排除した際の、説明力の変化と比較
 - ▶ 観察可能な変数の説明力と比較する

1.9 例: 学歴の効果

- 家庭環境の違い = U
 - ▶ 年齢と相関があると予想する

1.10 例: 学歴の効果

```
lm(earnings ~ education + age, CPSSW8) # genderの除外
```

```
Call:
lm(formula = earnings ~ education + age, data = CPSSW8)
```

```
Coefficients:
(Intercept)    education         age
   -11.1778      1.7064      0.1535
```

```
lm(earnings ~ education + gender, CPSSW8) # ageの除外
```

Call:

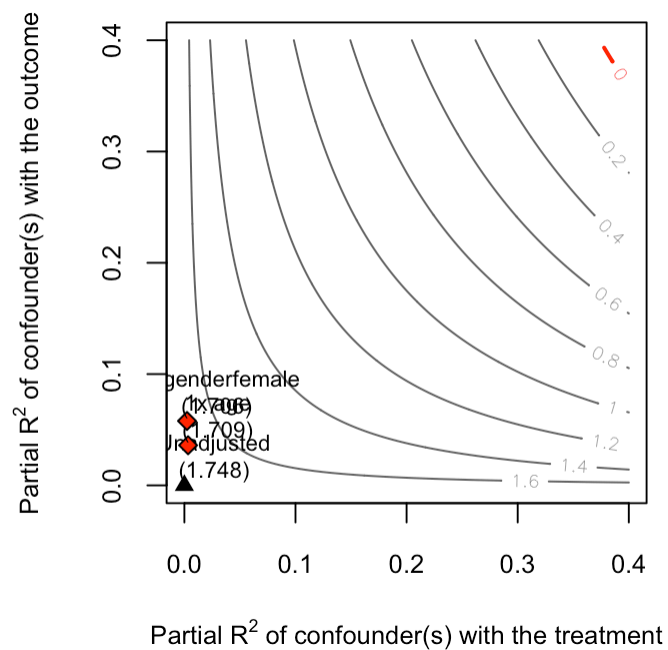
```
lm(formula = earnings ~ education + gender, data = CPSSW8)
```

Coefficients:

(Intercept)	education	genderfemale
-4.104	1.787	-4.188

1.11 例

```
sensemkr::ovb_contour_plot(  
  model,  
  treatment = "education",  
  benchmark_covariates = c("age", "genderfemale")  
)
```



- U が年齢と同程度に重要であったとしても、結果はあまり変化しない

1.12 Reference

Bibliography

- Cinelli, C., & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82(1), 39–67.
- Cinelli, C., Ferwerda, J., & Hazlett, C. (2024). *sensemkr*: Sensitivity analysis tools for OLS in R and Stata. *Observational Studies*, 10(2), 93–127.