第 14 回　回帰不連続デザイン（11）

村澤 康友

2020 年 7 月 21 日

今日のポイント

1. 一定のルールで処置の有無が決まる場合，処置確率に不連続性が生じるため，回帰式も不連続となる．回帰式の不連続性を利用して ATE を推定する手法を回帰不連続デザイン（RDD）という．
2. RDD は回帰式が不連続な点で条件付き ATE を局所的に推定する．重回帰モデルを仮定すれば OLS でも推定できる．
3. 処置確率が 0 から 1 にジャンプする RDD をシャープな RDD，それ以外をファジーな RDD という．
4. ファジーな RDD も回帰式が不連続な点で条件付き ATE を局所的に推定する．均一な処置効果を仮定すれば IV 法でも推定できる．

目次

1　回帰不連続デザイン

1.1　処置の割当ルール

無作為や自己選択でなく，一定のルールで処置の有無が決まる状況を考える．D を処置ダミー，X を共変量とする．X が基準値以上だと処置をするなら

$$D := [X \geq c]$$

ただし $[.]$ は中の命題が真なら 1，偽なら 0 を返す指示関数．この場合，処置群と対照群に X の値が等しい観測値は存在せず，マッチング法は使えない．

1.2　条件付き平均処置効果（ATE）

(Y_1^*, Y_0^*) を処置をする時としない時の潜在的な結果とする．$d = 0, 1$ について，Y_d^* の X 上へのノンパラメトリックな回帰モデルを仮定する．

$$\mathrm{E}(Y_d^*|X) = r_d(X)$$

$X = x$ のときの条件付き ATE は

$$\begin{aligned}
\mathrm{ATE}(x) &:= \mathrm{E}(Y_1^* - Y_0^*|X = x) \\
&= \mathrm{E}(Y_1^*|X = x) - \mathrm{E}(Y_0^*|X = x) \\
&= r_1(x) - r_0(x)
\end{aligned}$$

1.3　不連続な回帰式

$D := [X \geq c]$ なら観測される結果は

$$\begin{aligned}
Y &:= DY_1^* + (1 - D)Y_0^* \\
&= [X \geq c]Y_1^* + [X < c]Y_0^*
\end{aligned}$$

補題 1.

$$\mathrm{E}(Y|X) = [X \geq c] r_1(X) + [X < c] r_0(X)$$

証明.

$$\begin{aligned}\mathrm{E}(Y|X) &= \mathrm{E}([X \geq c] Y_1^* + [X < c] Y_0^* | X) \\ &= [X \geq c]\,\mathrm{E}(Y_1^*|X) + [X < c]\,\mathrm{E}(Y_0^*|X) \\ &= [X \geq c] r_1(X) + [X < c] r_0(X)\end{aligned}$$

□

注 1. すなわち $\mathrm{E}(Y|X)$ は $X = c$ で不連続（図 1）.

1.4　回帰不連続デザイン（RDD）（p. 254）

定義 1. 回帰式の不連続性を利用して ATE を推定する手法を**回帰不連続デザイン**（*Regression Discontinuity Design, RDD*）という.

定理 1. $r_0(.), r_1(.)$ が c で連続なら

$$\mathrm{ATE}(c) = \lim_{x \downarrow c} \mathrm{E}(Y|X = x) - \lim_{x \uparrow c} \mathrm{E}(Y|X = x)$$

証明.

$$\begin{aligned}&\lim_{x \downarrow c} \mathrm{E}(Y|X = x) - \lim_{x \uparrow c} \mathrm{E}(Y|X = x) \\ &= \lim_{x \downarrow c} r_1(x) - \lim_{x \uparrow c} r_0(x) \\ &= r_1(c) - r_0(c) \\ &= \mathrm{ATE}(c)\end{aligned}$$

□

注 2. 各項は $X = c$ 近傍の局所的な回帰モデルで推定できる.

1.5　重回帰モデル（p. 259）

次の重回帰モデルを仮定する.

$$\mathrm{E}(Y|D, X) = \alpha + \beta D + \gamma X + \delta D X$$

定理 2.

$$\mathrm{ATE}(c) = \beta + \delta c$$

証明. $D := [X \geq c]$ より

$$\begin{aligned}\lim_{x \downarrow c} \mathrm{E}(Y|X = x) &= \lim_{x \downarrow c} \mathrm{E}(Y|D = 1, X = x) \\ &= \lim_{x \downarrow c} (\alpha + \beta + \gamma x + \delta x) \\ &= \alpha + \beta + \gamma c + \delta c \\ \lim_{x \uparrow c} \mathrm{E}(Y|X = x) &= \lim_{x \uparrow c} \mathrm{E}(Y|D = 0, X = x) \\ &= \lim_{x \uparrow c} (\alpha + \gamma x) \\ &= \alpha + \gamma c\end{aligned}$$

第 1 式から第 2 式を引くと

$$\begin{aligned}\mathrm{ATE}(c) &= \alpha + \beta + \gamma c + \delta c - (\alpha + \gamma c) \\ &= \beta + \delta c\end{aligned}$$

□

注 3. 多項式回帰モデルにも拡張可能．ただし大域的な回帰モデルの定式化は誤りかもしれない.

2　ファジーな RDD

2.1　処置確率（p. 259）

$D := [X \geq c]$ なら

$$\Pr[D = 1|X] = [X \geq c]$$

すなわち処置確率は $X = c$ で 0 から 1 にジャンプする．より一般的に，処置確率が $X = c$ で p から q にジャンプする状況を考える．ただし $p, q \in [0, 1]$.

定義 2. 処置確率が 0 から 1 にジャンプする RDD を**シャープな** *RDD* という.

定義 3. シャープでない RDD を**ファジーな** *RDD* という.

例 1. X, Z を共変量とする．X, Z が同時に基準値以上だと処置をするなら

$$D := [X \geq c, Z \geq d]$$

Z が観測されないと

$$\begin{aligned}\Pr[D = 1|X] &= \Pr[X \geq c, Z \geq d|X] \\ &= \Pr[Z \geq d|X \geq c, X] \Pr[X \geq c|X] \\ &= \Pr[Z \geq d|X][X \geq c]\end{aligned}$$

すなわち処置確率は $X = c$ で 0 から $\Pr[Z \geq d|X]$ にジャンプする.

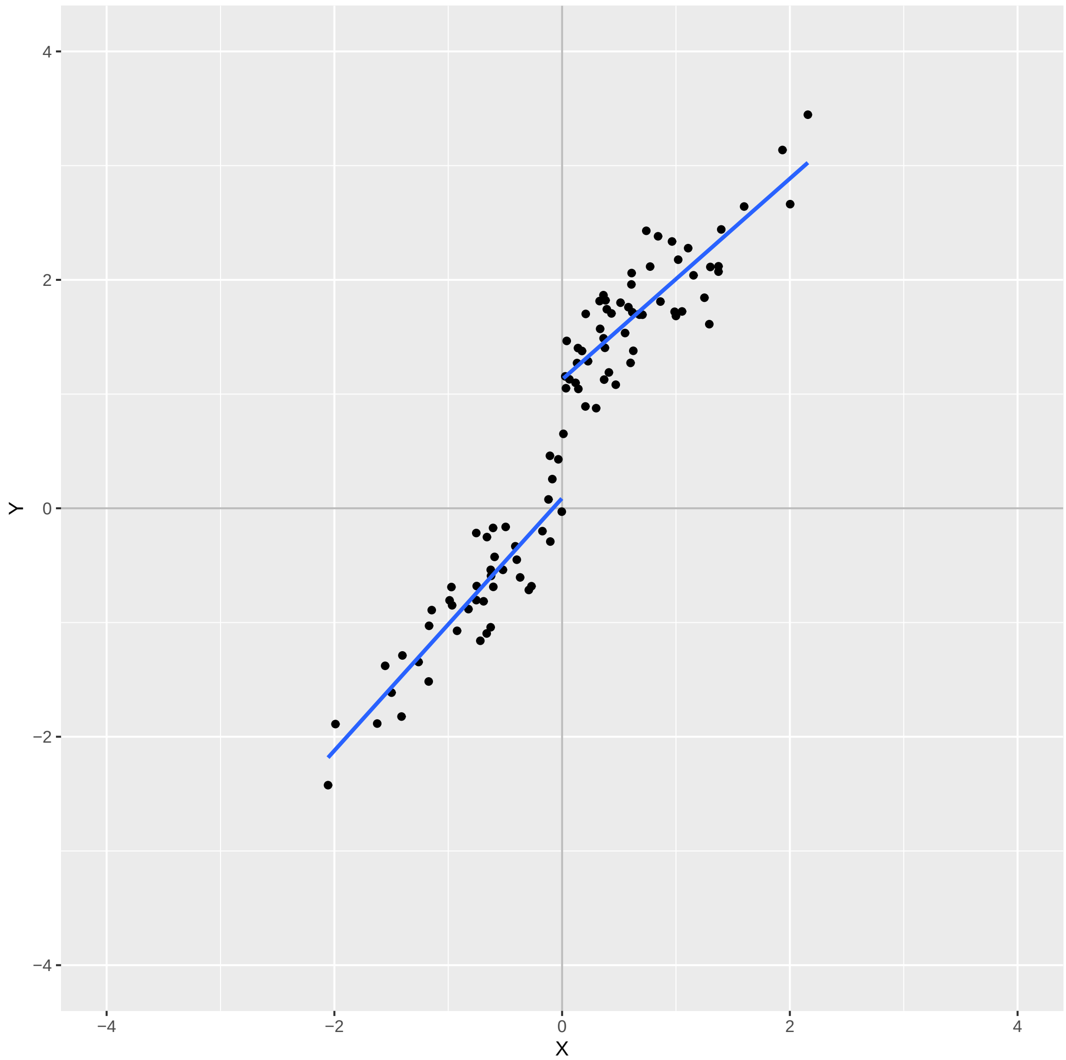

図 1　$X=0$ で不連続な回帰モデル

2.2　不連続な回帰式

$p(X) := \Pr[D=1|X]$ とする．$p(.)$ の c における右極限を $p(c+)$，左極限を $p(c-)$ と表す．すなわち

$$p(c+) := \lim_{x \downarrow c} p(x)$$
$$p(c-) := \lim_{x \uparrow c} p(x)$$

補題 2. X を所与として Y_1^*, Y_0^* が D と条件付き平均独立なら

$$\mathrm{E}(Y|X) = p(X)r_1(X) + (1-p(X))r_0(X)$$

証明．繰り返し期待値の法則より

$$\begin{aligned}
&\mathrm{E}(Y|X) \\
&= \mathrm{E}(DY_1^* + (1-D)Y_0^*|X) \\
&= \mathrm{E}(\mathrm{E}(DY_1^* + (1-D)Y_0^*|D,X)|X) \\
&= \mathrm{E}(D\,\mathrm{E}(Y_1^*|D,X) + (1-D)\,\mathrm{E}(Y_0^*|D,X)|X) \\
&= \mathrm{E}(D\,\mathrm{E}(Y_1^*|X) + (1-D)\,\mathrm{E}(Y_0^*|X)|X) \\
&= \mathrm{E}(D|X)\,\mathrm{E}(Y_1^*|X) + (1-\mathrm{E}(D|X))\,\mathrm{E}(Y_0^*|X) \\
&= \Pr[D=1|X]r_1(X) + (1-\Pr[D=1|X])r_0(X) \\
&= p(X)r_1(X) + (1-p(X))r_0(X)
\end{aligned}$$

□

注 4. $p(.)$ が c で不連続なら $\mathrm{E}(Y|X)$ も $X=c$ で

不連続．

2.3　ファジーな RDD（p. 259）

定理 3. X を所与として Y_1^*, Y_0^* が D と条件付き平均独立で，$r_0(.), r_1(.)$ が c で連続なら

$$\mathrm{ATE}(c) = \frac{\lim_{x\downarrow c} \mathrm{E}(Y|X=x) - \lim_{x\uparrow c} \mathrm{E}(Y|X=x)}{p(c+) - p(c-)}$$

証明. 補題より

$$\begin{aligned}
&\lim_{x\downarrow c} \mathrm{E}(Y|X=x) \\
&= \lim_{x\downarrow c}(p(x)r_1(x) + (1-p(x))r_0(x)) \\
&= p(c+)r_1(c) + (1-p(c+))r_0(c) \\
&= r_0(c) + p(c+)(r_1(c) - r_0(c)) \\
&\lim_{x\uparrow c} \mathrm{E}(Y|X=x) \\
&= \lim_{x\uparrow c}(p(x)r_1(x) + (1-p(x))r_0(x)) \\
&= p(c-)r_1(c) + (1-p(c-))r_0(c) \\
&= r_0(c) + p(c-)(r_1(c) - r_0(c))
\end{aligned}$$

第 1 式から第 2 式を引くと

$$\begin{aligned}
&\lim_{x\downarrow c} \mathrm{E}(Y|X=x) - \lim_{x\uparrow c} \mathrm{E}(Y|X=x) \\
&= (p(c+) - p(c-))(r_1(c) - r_0(c)) \\
&= (p(c+) - p(c-))\mathrm{ATE}(c)
\end{aligned}$$

整理すれば結果が得られる． □

注 5. 各項は $X = c$ 近傍の局所的な回帰モデルで推定できる．

2.4　均一な処置効果

均一な処置効果を仮定する．すなわち

$$\begin{aligned}
Y_0^* &= r(X) + U \\
Y_1^* &= Y_0^* + \beta \\
\mathrm{E}(U|X) &= 0
\end{aligned}$$

観測される結果は

$$\begin{aligned}
Y &:= DY_1^* + (1-D)Y_0^* \\
&= Y_0^* + (Y_1^* - Y_0^*)D \\
&= r(X) + \beta D + U
\end{aligned}$$

$r(X) := \alpha + \gamma X$ とすると

$$Y = \alpha + \beta D + \gamma X + U$$

$\mathrm{cov}(D, U) \neq 0$ だと OLS 推定量に偏りが生じる．

2.5　IV 推定（p. 260）

D はダミー変数なので $\Pr[D=1|X] = \mathrm{E}(D|X)$．$\mathrm{E}(D|X)$ が $X = c$ でジャンプするなら $\mathrm{E}(D|X)$ は $Z := [X \geq c]$ と相関する．

定理 4.

$$\begin{aligned}
\mathrm{E}(ZD) &\neq 0 \\
\mathrm{E}(ZU) &= 0
\end{aligned}$$

証明. $Z := [X \geq c]$ は X で一意に決まるので，繰り返し期待値の法則より

$$\begin{aligned}
\mathrm{E}(ZD) &= \mathrm{E}(\mathrm{E}(ZD|X)) \\
&= \mathrm{E}(Z\,\mathrm{E}(D|X)) \\
&= \mathrm{E}([X \geq c]\,\mathrm{E}(D|X)) \\
&\neq 0 \\
\mathrm{E}(ZU) &= \mathrm{E}(\mathrm{E}(ZU|X)) \\
&= \mathrm{E}(Z\,\mathrm{E}(U|X)) \\
&= 0
\end{aligned}$$

□

注 6. したがって Z を IV とした IV 法で均一な処置効果を推定できる．

注 7. D が 2 値変数でない場合にも拡張可能．

3　今日のキーワード

回帰不連続デザイン（RDD），シャープな RDD，ファジーな RDD

4　次回までの準備

提出　宿題 7，復習テスト 1–14（提出方法は追って連絡）

復習　教科書第 11 章，復習テスト 14

試験　持ち帰り試験（My KONAN で提出）