

末石計量 補足資料

前川 大空 *

2025 年 5 月 26 日

1 線形回帰と OLS

1.1 単回帰モデル

■p.3 『平均独立は独立性を示唆しない』 まず, 確率変数の独立性の仮定を確認する.

Def: 確率変数の独立性

確率変数 X, Y が $P(XY) = P(X)P(Y)$ ならば, X, Y は独立であるという.

平均独立の仮定より以下を得る:

$$\begin{aligned}\mathbb{E}[u_i] &\stackrel{\text{LIE}}{=} \mathbb{E}[\mathbb{E}[u_i|X_i]] \stackrel{\text{A2}}{=} \mathbb{E}[0] = 0 \\ \therefore \mathbb{E}[u_i] &= \mathbb{E}[u_i|X_i]\end{aligned}$$

また, 以下の関係式も成立する:

$$\begin{aligned}\implies \text{Cov}[u_i X_i] &= \mathbb{E}[u_i X_i] - \mathbb{E}[u_i]\mathbb{E}[X_i] \stackrel{\text{LIE}}{=} \mathbb{E}[X_i \mathbb{E}[u_i|X_i]] - \mathbb{E}[u_i]\mathbb{E}[X_i] \stackrel{\text{A2}}{=} 0 \\ \mathbb{E}[u_i X_i] &= \mathbb{E}[u_i]\mathbb{E}[X_i]\end{aligned}$$

これは独立性の必要条件に過ぎず, 誤差項と説明変数の**独立性を意味しない**. 仮定は平均が X に依存しないことを意味するのみで, **不均一分散等の可能性は排除しない**.

■p.3 単回帰モデルの仮定 1 単回帰モデルの仮定と誤差項の関係について考えたい.

単回帰モデルの仮定 1:

$(X_i, Y_i) \forall i \in \{1, \dots, n\}$ は i.i.d.

u_i の i.i.d. 性も確認する仮定として考えてよいのか? 特に独立性の定義と運用があやふやなので今一度確認したい. $(X_i, Y_i) \forall i \in \{1, \dots, n\}$ は独立. すなわち, 以下が全ての $i \neq j$ で成立していることを指す:

$$\begin{aligned}P((Y_i, X_i) \in A, (Y_j, X_j) \in B) &= P((Y_i, X_i) \in A)P((Y_j, X_j) \in B) \quad \forall A, B \subset \Omega \\ P_{X_i, Y_i, X_j, Y_j}(x_i, y_i, x_j, y_j) &= P_{X_i, Y_i}(x_i, y_i)P_{X_j, Y_j}(x_j, y_j) \quad \forall x_i, y_i \in \Omega\end{aligned}$$

基本的に下の記法を用いて独立性は議論する. 以下では独立を $(X_i, Y_i) \perp\!\!\!\perp (X_j, Y_j)$ とも表記する.

* 一橋大学経済学部 4 年, 五年一貫専修コース公共経済プログラム

■単回帰モデルにおける独立性 上記で挙げた疑問点に答えを与える. $(X_i, Y_i) \perp\!\!\!\perp (X_j, Y_j)$ から, $(X_i) \perp\!\!\!\perp (X_j)$, $(Y_i) \perp\!\!\!\perp (Y_j)$, $(X_i) \perp\!\!\!\perp (Y_j)$ 等は直ちに言えそうだが, 方針確認のため証明する.

Proof. 定義より以下が成立:

$$P_{X_i, Y_i, X_j, Y_j}(x_i, y_i, x_j, y_j) = P_{X_i, Y_i}(x_i, y_i) P_{X_j, Y_j}(x_j, y_j)$$

上の式と条件付期待値の定義より,

$$\begin{aligned} P_{X_i, Y_i, X_j, Y_j}(x_i, y_i, x_j, y_j) &= P_{X_i, Y_i}(x_i, y_i) P_{X_j, Y_j|X_i, Y_i}(x_j, y_j|x_i, y_i) \\ P_{X_j, Y_j|X_i, Y_i}(x_j, y_j|x_i, y_i) &= P_{X_j, Y_j}(x_j, y_j) \end{aligned}$$

条件付同時 pdf が無条件同時 pdf と一致する. この式自体も独立性の定義の一つ. 先ずは $(X_i) \perp\!\!\!\perp (X_j)$, 即ち

$$P_{X_j|X_i}(x_j|x_i) \stackrel{\text{def}}{=} \frac{P_{X_i, X_j}(x_i, x_j)}{P_{X_i}(x_i)} = \frac{P_{X_i}(x_i) P_{X_j}(x_j)}{P_{X_i}(x_i)} = P_{X_j}(x_j)$$

を証明したい. ここで,

$$\begin{aligned} P_{X_i, X_j}(x_i, x_j) &= \iint_{\mathbb{R}^2} P_{X_i, Y_i, X_j, Y_j}(x_i, y_i, x_j, y_j) dy_i dy_j \\ &\stackrel{\text{def}}{=} \iint_{\mathbb{R}^2} P_{X_i, Y_i}(x_i, y_i) P_{X_j, Y_j}(x_j, y_j) dy_i dy_j \\ &= \int_{\mathbb{R}} P_{X_i, Y_i}(x_i, y_i) dy_i \int_{\mathbb{R}} P_{X_j, Y_j}(x_j, y_j) dy_j = P_{X_i}(x_i) P_{X_j}(x_j) \end{aligned}$$

よって証明したい等式が成立. 他二つも同様に証明が可能. \square

次に, 上で示した $(X_i) \perp\!\!\!\perp (X_j)$, $(Y_i) \perp\!\!\!\perp (Y_j)$, $(X_i) \perp\!\!\!\perp (Y_j)$ を用いて, 誤差項の独立性 $(u_i) \perp\!\!\!\perp (u_j)$ を証明したい.

Proof. 単回帰モデルとその仮定:

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \text{ where } (X_i, Y_i) \stackrel{\text{i.i.d.}}{\sim} \forall i \in \{1, \dots, n\}$$

より, $Y_i - \beta_0 - \beta_1 X_i = u_i(X_i, Y_i) = u_i$. つまり, 示したいのは,

$$(u_i) \perp\!\!\!\perp (u_j), \text{ where } u_i = u_i(X_i, Y_i) = Y_i - \beta_0 - \beta_1 X_i$$

$$P_{u_i, u_j}(x_i, y_i, x_j, y_j) = P_{u_i}(x_i, y_i) P_{u_j}(x_j, y_j)$$

$u_i = Y_i - bX_i$ と $u_j = Y_j - bX_j$ との独立性を示すために, 各々の確率密度がどう計算されるか, それらが積の形に分解できるかを考える. $u_i = z$ の確率に注目する. 定義から $u_i = z \implies Y_i = bX_i + z$. $X_i = x$ のとき, $Y_i = bx + z$ となる確率は:

$$P_{u_i}(z) = \int_{\mathbb{R}} P_{X, Y}(x, bx + z) dx$$

という形で書ける. 同様に $u_j = w$ となる確率密度は:

$$P_{u_j}(w) = \int_{\mathbb{R}} P_{X, Y}(x', bx' + w) dx'$$

ここで u_i と u_j の同時密度を考えると:

$$\begin{aligned} P_{u_i, u_j}(z, w) &= \iint_{\mathbb{R}^2} P_{X_i, Y_i, X_j, Y_j}(x, bx + z, x', bx' + w) dx dx' \\ &\stackrel{\text{i.i.d.}}{=} \iint_{\mathbb{R}^2} P_{X_i, Y_i}(x, bx + z) P_{X_j, Y_j}(x', bx' + w) dx dx' \\ &= \int_{\mathbb{R}} P_{X, Y}(x, bx + z) dx \int_{\mathbb{R}} P_{X, Y}(x', bx' + w) dx' \stackrel{\text{def}}{=} P_{u_i}(z) P_{u_j}(w) \end{aligned}$$

よって目的の関係式を得た。 \square

よって, $(u_i) \perp (u_j)$ が確認できた。以上の議論より, 誤差項同士の独立性は仮定 1 のみで保証可能である。しかし, 同一分布性は先述の通り保証されず, 以下のように結論づけられる。

単回帰モデルの仮定下での誤差項の性質:

単回帰モデルの仮定の下で, 誤差項は独立だが i.i.d. ではない。

■p.4 識別と A2 観測可能な変数 (Y_i, X_i) の同時分布が分かれば, というのがミソ。識別の定義: 『観測されるデータの分布が既知の時, θ の値が一意に定まるならば, θ は識別されるという』に沿った記述であることを確認すること。より詳しく言えば, (X_i, Y_i) の同時分布が判明することで, $\mathbb{E}[Y_i|X_i]$ が X_i の関数として決定される。ここで (1.2) 式は X_i の一変数関数として考えられるため, β が一意に定まる (=識別される)。

■p.6 『漸近分散』の語の用法 自分の Note での用法と違うことに注意。状況によって柔軟に判断しよう。例えば, p.6 や p.9 で出てくる漸近分散は正確に言えば $\sqrt{n}(\hat{\beta}_1 - \beta_1)$ のものである。対して標準誤差という語は $\hat{\beta}_1$ そのものに用いられていることを注意しよう。

■p.6 $\sqrt{n}(\hat{\beta}_1 - \beta_1)$ の厳密分布 中級計量で与えられた説明は, 古典的仮定, 特に均一分散と正規性 (条件付きで良い) の下での議論。 $\sqrt{n}(\hat{\beta}_1 - \beta_1)$ を標準化したものは, 分散が観測可能な場合は標準正規分布に, 不可能な場合には自由度 $n - 2$ の t 分布に従っていた。この仮定を外した場合の議論は p.9 で少し展開されている。重要なのは漸近理論が適応可能で, この漸近分布はどの場合でも正規分布であることだ。

■p.8 『i.i.d. 仮定下で誤差項の無条件分散はすべての i について必ず等しくなる』

Proof. 以下の単回帰モデルと仮定のもとで議論を行う:

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \text{ where } (X_i, Y_i) \text{ は RVs, } \stackrel{\text{i.i.d.}}{\sim} \forall i \in \{1, \dots, n\}$$

$\text{Var}(u_i)$ が i に関し一定であると示す。 $\text{Var}(Y_i) = \sigma_y^2, \text{Var}(X_i) = \sigma_x^2$ とすると: $\sigma_y^2 = \text{Var}(\beta_1 X_i + u_i) = \beta_1^2 \sigma_x^2 + \text{Var}(u_i) + 2\beta_1 \text{Cov}(X_i, u_i), (\text{const.}) = \text{Var}(u_i) + 2\beta_1 \text{Cov}(X_i, u_i) \stackrel{\text{A2}}{=} \text{Var}(u_i)$. \square

不均一分散の定義:

単回帰モデルの仮定の下では無条件分散の不均一分散は実現せず, p.7 の形で定義を行う必要がある。

■p.9 不均一分散下での $\hat{\beta}_1$ の標準誤差導出 ここでは OLSE の分散を求めて, その標本対応から導いた標準誤差が古典的仮定のもとでの標準誤差とは異なることを示す。単回帰モデルとその仮定:

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \text{ where } (X_i, Y_i) \stackrel{\text{i.i.d.}}{\sim} \forall i \in \{1, \dots, n\}$$

OLS 推定量は

$$(\hat{\beta}_0, \hat{\beta}_1) = (\bar{Y} - \hat{\beta}_1 \bar{X}, \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}) = (\beta_0 + (\beta_1 - \hat{\beta}_1) \bar{X} + \bar{u}, \beta_1 + \frac{\sum_{i=1}^n (X_i - \bar{X})u_i}{\sum_{i=1}^n (X_i - \bar{X})^2})$$

まず不偏性を示そう。

Proof.

$$\begin{aligned} \mathbb{E}[\hat{\beta}_1] &= \beta_1 + \mathbb{E}\left[\frac{\sum_{i=1}^n (X_i - \bar{X})u_i}{\sum_{i=1}^n (X_i - \bar{X})^2}\right] \\ &\stackrel{\text{LIE}}{=} \beta_1 + \mathbb{E}\left[\mathbb{E}\left[\frac{\sum_{i=1}^n (X_i - \bar{X})u_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \middle| \mathbf{X}\right]\right] \\ &= \beta_1 + \mathbb{E}\left[\frac{\sum_{i=1}^n (X_i - \bar{X})\mathbb{E}[u_i | \mathbf{X}]}{\sum_{i=1}^n (X_i - \bar{X})^2}\right] \\ &\stackrel{\text{i.i.d.}}{=} \beta_1 + \mathbb{E}\left[\frac{\sum_{i=1}^n (X_i - \bar{X})\mathbb{E}[u_i | X_i]}{\sum_{i=1}^n (X_i - \bar{X})^2}\right] \stackrel{\text{A2}}{=} \beta_1 \end{aligned}$$

$\hat{\beta}_0$ も LIE により不偏性を確認できる。 □

$$\text{Var}(\hat{\beta}_1) = \mathbb{E}[(\hat{\beta}_1 - \beta_1)^2] = \mathbb{E}\left[\left(\frac{\sum_{i=1}^n (X_i - \bar{X})u_i}{\sum_{i=1}^n (X_i - \bar{X})^2}\right)^2\right] \stackrel{\text{LIE}}{=} \mathbb{E}\left[\frac{\mathbb{E}[(\sum_{i=1}^n (X_i - \bar{X})u_i)^2 | \mathbf{X}]}{[\sum_{i=1}^n (X_i - \bar{X})^2]^2}\right]. \text{ 期待値内の分母を考えよう.}$$

$$(\text{denominator}) = \mathbb{E}\left[\left(\sum_{i=1}^n (X_i - \bar{X})u_i\right)^2 \middle| \mathbf{X}\right] = \mathbb{E}[S^2 | \mathbf{X}] \text{ where } S = \sum_{i=1}^n (X_i - \bar{X})u_i$$

$$\mathbb{E}[S^2 | \mathbf{X}] = \sum_{i=1}^n (X_i - \bar{X})^2 \mathbb{E}[u_i^2 | \mathbf{X}] + 2 \sum_{i < j} (X_i - \bar{X})(X_j - \bar{X}) \mathbb{E}[u_i u_j | \mathbf{X}].$$

先述の議論より, A1 より誤差項同士が独立, つまり相関がない (\because 独立性の必要条件) ので:

$$\mathbb{E}[u_i u_j | \mathbf{X}] = \mathbb{E}[u_i | \mathbf{X}] \mathbb{E}[u_j | \mathbf{X}] = 0, \therefore \mathbb{E}[S^2 | \mathbf{X}] = \sum_{i=1}^n (X_i - \bar{X})^2 \mathbb{E}[u_i^2 | \mathbf{X}] = \sigma_i^2 \forall i \neq j$$

$$\text{Var}(\hat{\beta}_1) = \mathbb{E}\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2 \mathbb{E}[u_i^2 | \mathbf{X}]}{[\sum_{i=1}^n (X_i - \bar{X})^2]^2}\right]$$

$$\text{Var}(\hat{\beta}_1 | \mathbf{X}) \stackrel{\text{LIE}}{=} \frac{\sum_{i=1}^n (X_i - \bar{X})^2 \mathbb{E}[u_i^2 | \mathbf{X}]}{[\sum_{i=1}^n (X_i - \bar{X})^2]^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2 \sigma_i^2}{[\sum_{i=1}^n (X_i - \bar{X})^2]^2}$$

更に均一分散の仮定を置けば:

$$\text{Var}(u_i | X_i) = \mathbb{E}[u_i^2 | \mathbf{X}] = \sigma^2 \forall i, \therefore \mathbb{E}[S^2 | \mathbf{X}] = \sigma^2 \sum_{i=1}^n (X_i - \bar{X})^2$$

$$\text{Var}(\hat{\beta}_1) = \sigma^2 \mathbb{E}\left[\frac{1}{\sum_{i=1}^n (X_i - \bar{X})^2}\right], \text{Var}(\hat{\beta}_1 | \mathbf{X}) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

古典的仮定下での OLSE の条件付分散が得られ, これは本文の記述と一致している. だが実際は $\text{Var}(u_i | X_i) = \sigma(X_i)$ と X_i に依存することがしばしばで, この場合古典的仮定下での標準誤差は推定には使えない.

■『単回帰モデルの仮定』下での漸近分散 条件付分散の確率収束先, 漸近分散の振る舞いを見よう.

Proof. 漸近分布という語は標準化を暗黙に要求していたことに注意して:

$$\begin{aligned}
 V = \text{Var}(\sqrt{n}(\hat{\beta}_1 - \beta_1) | \mathbf{X}) &= \frac{n \sum_{i=1}^n [(X_i - \bar{X})^2 \sigma_i^2]}{[\sum_{i=1}^n (X_i - \bar{X})^2]^2} \\
 &\stackrel{\text{heterogeneity}}{=} \frac{\frac{1}{n} \sum_{i=1}^n [(X_i - \bar{X})^2 \sigma^2(X_i)]}{[\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2]^2} \\
 &\stackrel{p}{\rightarrow} \frac{\mathbb{E}[(X_i - \mathbb{E}[X])^2 \sigma^2(X_i)]}{\sigma_X^4} \\
 &= \frac{\mathbb{E}[(X_i - \mathbb{E}[X])^2 \mathbb{E}[u_i^2 | \mathbf{X}]]}{\sigma_X^4} \\
 &\stackrel{\text{LIE}}{=} \frac{\mathbb{E}[(X_i - \mathbb{E}[X])^2 u_i^2]}{\sigma_X^4} = \frac{\text{Var}((X_i - \mathbb{E}[X])u_i)}{\text{Var}(X_i)^2}
 \end{aligned}$$

ここでは連続写像定理, 4 次モーメントまでの有限性を用いた. 教科書の記述と合致する. \square

■誤差項に関する期待値計算の性質 上記の導出において用いた誤差項に関する期待値計算の性質の正統性を確認しておこう. まず, 不偏性の証明で使った $\mathbb{E}[u_i | \mathbf{X}] \stackrel{\text{i.i.d.}}{=} \mathbb{E}[u_i | X_i]$ を確認する. 次に, $\mathbb{E}[u_i u_j | \mathbf{X}] = \mathbb{E}[u_i | X_i] \mathbb{E}[u_j | X_j]$ と, この結果に必要な仮定を clarify しておこう.

Proof. $\mathbb{E}[u_i | \mathbf{X}] = \mathbb{E}[u_i | X_i]$ を示したい. つまり以下が成り立つことを確かめたい:

$$\begin{aligned}
 \mathbb{E}[u_i | \mathbf{X}] &= \int u_i f_{u_i | \mathbf{X}}(u_i | \mathbf{X}) du_i \\
 &\stackrel{(\dagger)}{=} \int u_i f_{u_i | X_i}(u_i | X_i) du_i = \mathbb{E}[u_i | X_i].
 \end{aligned}$$

ここで (†) を示したい. 単回帰モデルの仮定

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \quad (X_i, Y_i) \stackrel{\text{i.i.d.}}{\sim}$$

より $u_i = u_i(X_i, Y_i) = Y_i - \beta_0 - \beta_1 X_i$ は X_i と Y_i で決定され, \mathbf{X}_{-i} は一切関与しない. よって,

$$f_{u_i | \mathbf{X}}(u_i | \mathbf{X}) = f_{u_i | X_i, \mathbf{X}_{-i}}(u_i | X_i, \mathbf{X}_{-i}) = f_{u_i | X_i}(u_i | X_i),$$

が成り立つため, (†) も示され題意が満たされた. \square

Proof. 次に, $\mathbb{E}[u_i u_j | \mathbf{X}] = \mathbb{E}[u_i | X_i] \mathbb{E}[u_j | X_j]$, $i \neq j$ を検証しよう. 下を示すのが目標だ:

$$\begin{aligned}
 \mathbb{E}[u_i u_j | \mathbf{X}] &= \iint u_i u_j f_{u_i, u_j | \mathbf{X}}(u_i, u_j | \mathbf{X}) du_i du_j \\
 &\stackrel{(\ddagger)}{=} \iint u_i u_j f_{u_i | X_i}(u_i | X_i) f_{u_j | X_j}(u_j | X_j) du_i du_j \\
 &= \left(\int u_i f_{u_i | X_i}(u_i | X_i) du_i \right) \left(\int u_j f_{u_j | X_j}(u_j | X_j) du_j \right) = \mathbb{E}[u_i | X_i] \mathbb{E}[u_j | X_j].
 \end{aligned}$$

ここで (‡) を示したい. 先に示した誤差項の独立性と, (X_i, Y_i) の i.i.d. 性から

$$f_{u_i, u_j | \mathbf{X}} = f_{u_i | X_i} f_{u_j | X_j}.$$

によって (‡) が成立し, したがって題意は満たされた. \square

■連続写像定理の正統性検証 次に、漸近分布の導出で用いた連続写像定理 (CMT) の正統性を検証しよう。

Thm: 連続写像定理 (Continuous Mapping Theorem)

確率収束 $X_n \xrightarrow{p} X$ を満たす確率変数列 $\{X_n\}$ と、関数 $g: \mathbb{R}^k \rightarrow \mathbb{R}$ が点 x の近傍で連続であるとき、

$$g(X_n) \xrightarrow{p} g(X) \quad (\text{ただし } X_n \xrightarrow{p} X).$$

では CMT を確認したところで、目的の式変形の正統性を確認しよう。なお、分母における分散の確率収束については以前検証したため、ここでは確認せずに確率収束を認める。

Proof. 標準化を暗黙に含むことに注意して、以下の分子・分母それぞれの確率収束を示す

$$V = \text{Var}(\sqrt{n}(\hat{\beta}_1 - \beta_1) | \mathbf{X}) = \frac{\frac{1}{n} \sum_{i=1}^n [(X_i - \bar{X})^2 \sigma^2(X_i)]}{[\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2]^2}.$$

$$(\text{denominator}) = A_n^2 = [\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2]^2 \xrightarrow{p} (\text{Var}(X_i))^2 = \sigma_X^4 \quad (\cdot: \text{LLN, CMT})$$

$$(\text{numerator}) = B_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \sigma^2(X_i) = \frac{1}{n} \sum_{i=1}^n f(X_i, \bar{X}),$$

$$\text{where } f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, m) = (x - m)^2 \sigma^2(x).$$

ここで $\sigma(x)$ の連続性を仮定すれば f は連続関数である。各 i について、 $(X_i, \bar{X}) \xrightarrow{p} (X_i, \mathbb{E}[X])$ である（第一成分は定常に同じ分布、第二成分は LLN による）。CMT より

$$f(X_i, \bar{X}) \xrightarrow{p} f(X_i, \mathbb{E}[X]) = (X_i - \mathbb{E}[X])^2 \sigma^2(X_i).$$

さらに、 X_i の i.i.d. 性と 4 次モーメントまでの有限性の下で LLN を再適用すると、

$$\frac{1}{n} \sum_{i=1}^n f(X_i, \bar{X}) \xrightarrow{p} \mathbb{E}[f(X_i, \mathbb{E}[X])] = \mathbb{E}[(X_i - \mathbb{E}[X])^2 \sigma^2(X_i)].$$

$$B_n \xrightarrow{p} \mathbb{E}[(X_i - \mathbb{E}[X])^2 \sigma^2(X_i)]$$

これらを組み合わせ、LLN, CMT, LIE を用いて目的の確率収束先を得た。□

1.2 重回帰モデル

■p.15 完全な多重共線性 ある説明変数が他の説明変数に線形従属である状態のことを指す。『R による実証分析』勉強会の 11, 12 章の資料でも散々議論したのでそちらも参照してほしい。今回は OLS 推定量の導出に注目した議論を、注 7 に絡めてしてみたい。（線形）回帰モデルの OLSE は以下の問題を解き与えられる：

$$\text{重回帰モデル: } \min_{\mathbf{b}} \sum_i [Y_i - \mathbb{E}[Y_i | \mathbf{X}_i]]^2 = \min_{\mathbf{b}} \sum_i [Y_i - \mathbf{X}_i \mathbf{b}]^2$$

$$\text{単回帰モデル: } \min_{b_0, b_1} \sum_i [Y_i - \mathbb{E}[Y_i | X_i]]^2 = \min_{b_0, b_1} \sum_i [Y_i - b_0 - b_1 X_i]^2$$

説明変数行列 \mathbf{X} の第一列 (それゆえ行ベクトル \mathbf{X}_i の第一成分も) の要素は全て 1 であることに注意せよ. 単回帰モデルは説明変数が定数項と X の 2 つのみの重回帰モデルと言える. この最小化問題で重要なのは, 解が一意に定まるかである. 注 7 の通り, 変数 X が一定になる場合, つまり:

$$b_0 + b_1 X_i = b_0 \cdot 1 + b_1 \cdot c = \hat{y} \text{ (constant)}$$

のとき, 我々の目的は先述の損失関数を最小化することにあつたが, 一定の \hat{y} に対して我々が選択できるのは (b_0, b_1) の二つである. これは不定方程式であり, \mathbf{b} の選び方は無数に存在するため識別が不可能である. さらにこの場合, 説明変数ベクトル X は定数項ベクトル $\mathbf{1}$ と完全な多重共線性を持つことが分かる. つまり:

$$X = \begin{pmatrix} c \\ \vdots \\ c \end{pmatrix} = c \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = c \cdot \mathbf{1}$$

以上より, 我々がよく知る OLSE を得るためには, 以下のように条件付けをするといことが分かった:

単回帰モデルの OLS 推定量

説明変数 X が一定でないとき (= 定数項との間に完全な多重共線性がないとき),

$$(\hat{\beta}_0, \hat{\beta}_1) = (\bar{Y} - \hat{\beta}_1 \bar{X}, \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}).$$

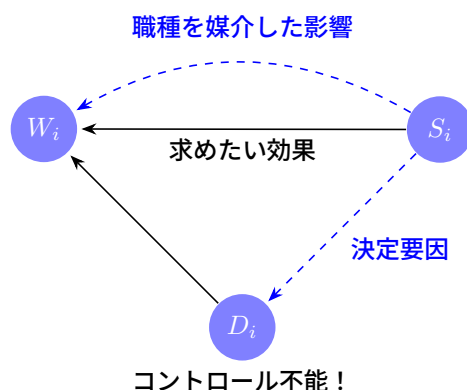
同様の議論が重回帰モデルの議論でも成り立つ. ^{*1} 末石計量では行列表記が 4 章以降に回されるので, ここでは結果だけ示しておこう. 導出は Econometrics Note で行っており, 記法もこれに従う. ^{*2}

重回帰モデルの OLS 推定量

説明変数間に完全な多重共線性がないとき,

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}.$$

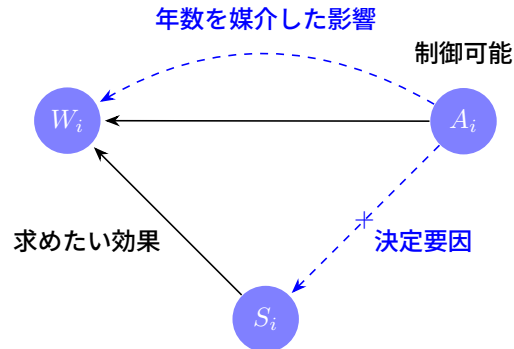
■p.16 不適切なコントロール変数 本文にあげられた例を図示してみよう.



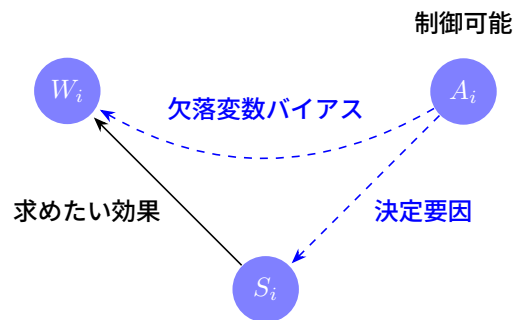
^{*1} 繰り返しになるが, Dummy Variable Trap の議論はまさに多重共線性について言及したものである.

^{*2} OLSE に使用される行列 $\mathbf{X}^T \mathbf{X}$ に逆行列が存在するのは列フルランク行列のみであるため, $\text{rank}(\mathbf{X}) = \text{rank}(\mathbf{X}^T \mathbf{X})$ で, 列ベクトルが線形独立でない (完全な多重共線性が存在する) \Leftrightarrow 正則ではない \Leftrightarrow 逆行列が存在しない. このように OLS 推定量の導出に用いられる公式は完全な多重共線性の下では機能しないが, そもそもモデル仮定の時点でこのようなケースは弾かれるのである.

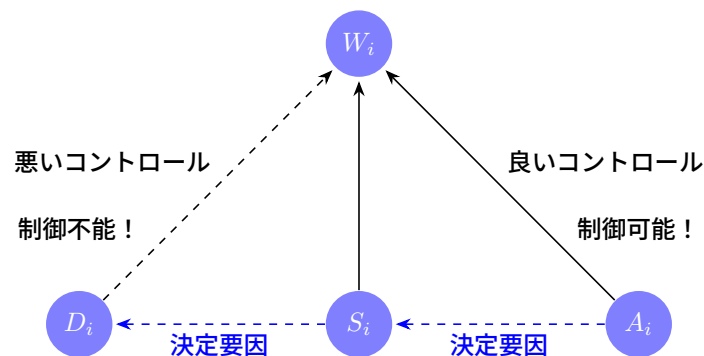
能力がコントロール変数として使えるのが何故か、職種との差異にも一考の余地がある。この例では能力がコントロール可能であるため、能力から教育年数への影響をうまくつぶすことが出来る。



欠落変数バイアスは以下のような図示が可能だ。



関係を整理するとこの通り。興味のある説明変数の決定要因になるものはコントロール変数として利用できるが、興味のある説明変数が決定要因になるものは利用できない (bad control)。



■p.17 不完全な多重共線性 Econometrics Note では冗長なコントロール変数の問題と絡めて、説明変数の相関関係が大きいときの問題を指摘したが、本文の記述の通りこれは一般的にも成り立つ。分散が大きくなってしまうものの、明確な対処法は存在しないため気にしすぎるものではない。

1.3 線形射影

■p.19 不完全な多重共線性 (未解決) 最適化問題の建付けとして, 線形射影は期待値の最小化である一方で, 回帰モデルではわざわざ期待値の (おそらく) 標本対応を使って議論を始めているのは何故だろうか. 一致推定できる, との結論に帰着するため?

■p.20 標本対応 標本対応は一致性により正当化される. Econometrics Note の 4 ページを確認すること.

■p.20 線形モデルの有用性 線形射影は Y_i のみならず, $\mathbb{E}[Y_i|\mathbf{X}_i]$ の線形最良近似である. 一つのデータのみならず, 分布についての情報を与えてくれる点で線形モデルは有用だと言える.

付録 A 計算上のテクニック

■畳み込み積分 (Convolution) 畳み込みとは, 和が一定となるようなものをかけて足し合わせるという操作のことである. これを積分で利用したのが畳み込み積分である.

$$h(x) = \int_{-\infty}^{\infty} f(t)g(x-t)dt$$

右辺は x の関数. この $h(x)$ を $f(x)$ と $g(x)$ の畳み込み積分 (あるいは単に畳み込み) などと呼び, $h(x) = f(x) * g(x)$ などと表記する. 確率変数 X と Y の確率分布が分かっているときに, $Z = X + Y$ の確率分布は畳み込みで与えられる. つまり:

$$h(z) = \int_{-\infty}^{\infty} f(x)g(z-x)dx$$

と計算できる. ここで, h, f, g はそれぞれ順に連続確率変数 $Z = X + Y, X, Y$ の確率密度関数である.

参考文献

- [1] 末石 直也 (2015), 計量経済学—マイクロデータ分析へのいざない, 第 1 版, 日本評論社