

# 伝統的な重力方程式の推定

田中 鮎夢

2024-05-12

## Table of Contents

貿易データ .....	1
対数変換 .....	2
推定 .....	3
推定結果の解釈 .....	4
連続変数 .....	4
ダミー変数 .....	4
推定結果の整形と表示1 with modelsummary .....	5
推定結果のLaTeX形式出力 .....	7
係数プロット with modelplot .....	7
推定結果の整形と表示2 with stargazer .....	8
参考文献 .....	9

## 貿易データ


本ページでは、重力方程式の基本的な推定方法を説明する。日本語の文献としては、重力方程式について、伊藤・田中『現実からまなぶ国際経済学』や田中「国際貿易と重力の意外な関係: 重力方程式の基本」に説明がある。

[gravity.xlsx](#) データには、2005 年の 1 年間の世界の 2 国間の貿易額が収録されている。

R でデータを読み込む。

```
library(readxl)
gravity <- read_excel("gravity.xlsx")
head(gravity)

## # A tibble: 6 × 20
##   importer exporter  year imports gdp_exporter gdp_importer join_exporter
##   <chr>    <chr>   <dbl>   <dbl>      <dbl>      <dbl>      <dbl>
## 1 AFG      ABW     2005     0         NA      6814753581      NA
```

```
## 2 AGO      ABW      2005      0      NA  30632364954      NA
## 3 ALB      ABW      2005      0      NA  8376483740      NA
## 4 ANT      ABW      2005  4335.      NA      NA      NA
## 5 ARE      ABW      2005      0.951      NA 133000000000      NA
## 6 ARG      ABW      2005      0.658      NA 183000000000      NA
## #  13 more variables: join_importer <dbl>, exporternum <dbl>,
## #   importernum <dbl>, contig <dbl>, comlang_off <dbl>, colony <dbl>,
## #   dist <dbl>, REPlandlocked <dbl>, PARTlandlocked <dbl>, religion <dbl>,
## #   onein <dbl>, bothin <dbl>, nonein <dbl>
```

オンラインから直接データを読み込みたい場合は、パッケージ `gdata` をインストールした上で、以下のコードを用いる。

```
# install.packages("gdata")
library(gdata)
library(readxl)

download.file("https://ayumu-
tanaka.github.io/teaching/gravity.xlsx", "gravity.xlsx", mode="wb")
gravity <- read_excel("gravity.xlsx")
```

変数 `importer` は輸入国の ISO コードと呼ばれる

3桁の国コードである。同様に、`exporter` は輸出国の ISO

コードである。例えば、日本の ISO コードは `JPN`、アメリカの ISO コードは `USA`

である。また、ISO コード (`exporter`,

`importer`) の他に、輸入額 (`imports`)、輸出国の GDP (`gdp_exporter`)、輸入国の GDP (`gdp_importer`)、輸出国と輸入国の間の距離 (`dist`) が含まれている。

## 対数変換

R で次のような重力方程式を推定することを考える。

$$\text{輸入額} = \text{定数} \times \frac{\text{輸出国のGDP}^\alpha \times \text{輸出国のGDP}^\beta}{\text{輸出国と輸入国の間の距離}^\gamma} \times e^{\delta \times \text{言語の共通性}}$$

この式は非線形のためコンピューターの能力上、推定が難しい。そのため、従来は両辺の対数をとって、線形にされてきた。線形にされた式の推定は最小二乗法で比較的簡単に推定できる。そこでまず、回帰分析に用いる変数の対数を取る。R のコードは次の通りである。

```
gravity$lnimports <- log(gravity$imports)
gravity$lngdp_exporter <- log(gravity$gdp_exporter)
gravity$lngdp_importer <- log(gravity$gdp_importer)
gravity$lnldist <- log(gravity$dist)
```

ここで、0 より大きい値しか対数値にできないことに注意が必要である。

そのため、輸入額が 0 より大きい値のサンプルを作成しておく。

```
gravity2 <- subset(gravity, gravity$imports>0)
```

## 推定

通常、貿易データは不均一分散の性質を持つので、デフォルトの`lm`で推定するのではなく、`lm_robust`で推定する。そのために、Blair et al. (2018) が開発したパッケージ`estimatr`をインストールする。

```
install.packages("estimatr")
```

パッケージ`estimatr`の使い方は、[公式ページ](#)で確認できる。

不均一分散頑健な標準誤差を計算するよう、`lm_robust`で重力方程式を推定するコードは以下の通りである。`summary(ols)`により回帰分析の結果が出力される。

```
library(estimatr)
ols <- lm_robust(limports ~
  lgdp_exporter + lgdp_importer + ldist + comlang_off,
  data = gravity2)

summary(ols)

##
## Call:
## lm_robust(formula = limports ~ lgdp_exporter + lgdp_importer +
##          ldist + comlang_off, data = gravity2)
##
## Standard error type: HC2
##
## Coefficients:
##              Estimate Std. Error t value   Pr(>|t|) CI Lower CI Upper
DF
## (Intercept)   -33.751    0.352954  -95.63 0.000e+00 -34.4433 -33.0596
19973
## lgdp_exporter    1.226    0.007648  160.29 0.000e+00  1.2108  1.2408
19973
## lgdp_importer    0.951    0.007571  125.61 0.000e+00  0.9361  0.9658
19973
## ldist          -1.374    0.020435  -67.24 0.000e+00 -1.4141 -1.3340
19973
## comlang_off      1.293    0.050422   25.65 8.613e-143  1.1945  1.3922
19973
##
## Multiple R-squared:  0.6424 ,    Adjusted R-squared:  0.6423
## F-statistic:  9877 on 4 and 19973 DF,  p-value: < 2.2e-16
```

なお、回帰分析は、以下のコードでも実行できる。

```
# ols<-lm_robust(gravity2$limports ~ gravity2$lgdp_exporter +  
gravity2$lgdp_importer + gravity2$ldist + gravity2$comlang_off)
```

## 推定結果の解釈

### 連続変数

推定結果において、たとえば、輸出国のGDP (lgdp\_exporter) の推定係数は  $\hat{\alpha} = 1.226$  になっている。これは、以下のような関係にあると言える。

$$\frac{\partial \ln \widehat{\text{輸入額}}}{\partial \ln \text{輸出国のGDP}} = \hat{\alpha} = 1.226$$

また、対数微分の公式

$$\frac{d \ln x}{dx} = \frac{1}{x}$$

から、 $d \ln x = \frac{dx}{x}$  となるので、

$d \ln x$  は変化率を表している。この関係を使って上の式を書き直すと、近似的に以下が成り立つ。

$$\frac{\text{輸入額の変化率}}{\text{輸出国のGDPの変化率}} = \hat{\alpha} = 1.226$$

この式は、弾性値 (=Yの変化率/Xの変化率) の形になっている。そのため、輸出国のGDPが1%大きくなれば、平均的に輸入額が1.226%大きくなる傾向にあると表現できる。

### ダミー変数

また、言語の共通性 (comlang\_off) の推定係数は  $\hat{\delta} = 1.293$  である。もともと、言語の共通性以外の他の項を  $B$  としてまとめると、重力方程式の定式化から以下のような関係にある。

$$\text{輸入額} = B \times e^{\delta \times \text{言語の共通性}}$$

そのため、言語が共通の場合は、

$$\widehat{\text{輸入額}} = B \times e^{1.293 \times 1} = B \times e^{1.293}$$

そのため、言語が共通ではない場合は、

$$\widehat{\text{輸入額}} = B \times e^{1.293 \times 0} = B \times e^0 = B$$

となる。言語が共通な場合は、

$$\frac{B \times e^{1.293}}{B} = e^{1.293} \cong 3.64$$

より、平均的に3.64倍貿易額が大きい傾向にあることが分かる。

なお、以下のように計算している。

```
b <- exp(1.293)
b
## [1] 3.643701
```

## 推定結果の整形と表示1 with modelsummary

### Regression Tables with

`estimatr`で説明されているように、`stargazer`ではなく、`Texreg`や`modelsummary`を用いた方が簡単に、`lm_robust`を用いた不均一分散頑健な推定結果を表示できる。

ここでは、Arel-Bundock (2022)

が開発した`modelsummary`を用いて、推定結果を表示する。

```
#install.packages("modelsummary")
library(modelsummary)
#(1) lmを使用して通常の回帰分析を実行し、ols1に保存。
ols1 <- lm(limports ~ lgdp_exporter
          + lgdp_importer + ldist + comlang_off,
          data = gravity2)

#(2)
lm_robustを使用して回帰分析を実行し、不均一分散頑健な標準誤差を計算し、ols2に保存。
ols2 <- lm_robust(limports ~ lgdp_exporter
                  + lgdp_importer + ldist + comlang_off,
                  data = gravity2)

modelsummary(list(ols1, ols2), stars = TRUE)
```

(Intercept)	-33.751*** (0.350)	-33.751*** (0.353)
lgdp_exporter	1.226*** (0.008)	1.226*** (0.008)
lgdp_importer	0.951*** (0.008)	0.951*** (0.008)
ldist	-1.374*** (0.022)	-1.374*** (0.020)

comlang_off	1.293*** (0.051)	1.293*** (0.050)
Num.Obs.	19978	19978
R2	0.642	0.642
R2 Adj.	0.642	0.642
AIC	93181.8	93181.8
BIC	93229.2	93229.2
Log.Lik.	-46584.884	
RMSE	2.49	2.49

- p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

より手の込んだ推定結果表も作成できる。参考: [modelsummary: regression tables](#)

```
modelsummary(
  list(ols1, ols2),
  estimate = c("estimate",
               "{estimate}{stars}"),
  coef_omit = "Intercept",
  gof_omit = 'DF|Deviance|AIC|BIC|RMSE',
  notes = list('* p<0.1; ** p<0.05; *** p<0.01')
)
```

lgdp_exporter	1.226 (0.008)	1.226*** (0.008)
lgdp_importer	0.951 (0.008)	0.951*** (0.008)
ldist	-1.374 (0.022)	-1.374*** (0.020)
comlang_off	1.293 (0.051)	1.293*** (0.050)
Num.Obs.	19978	19978
R2	0.642	0.642
R2 Adj.	0.642	0.642
Log.Lik.	-46584.884	

- p<0.1; \*\* p<0.05; \*\*\* p<0.01

## 推定結果のLaTeX形式出力

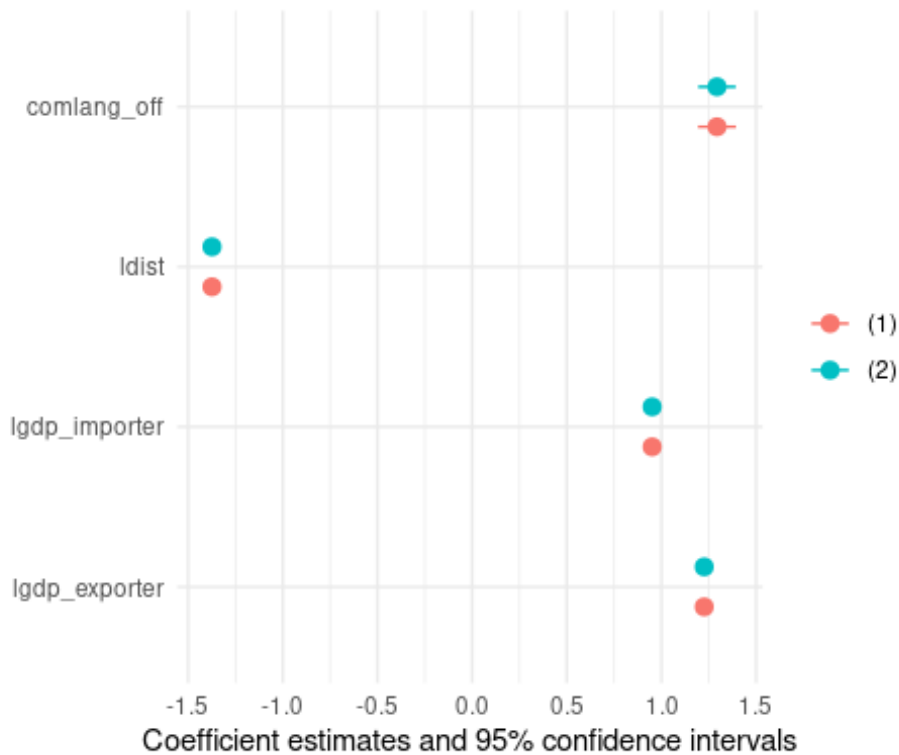
```
modelsummary(list(ols1, ols2), output = "table.tex")
```

LaTeXのpreambleに以下のコード記載することが必要。

```
\usepackage{booktabs}
\usepackage{siunitx}
\newcolumntype{d}{S[
  input-open-uncertainty=,
  input-close-uncertainty=,
  parse-numbers = false,
  table-align-text-pre=false,
  table-align-text-post=false
]}
```

## 係数プロット with modelplot

```
library(ggplot2)
modelplot(list(ols1, ols2), coef_omit = "Intercept")
```



```
ggsave("coefplot.png")
```





#(3) 通常のOLSの標準誤差は以下で表示できる。

```
stargazer(ols1, type="text")

##
## =====
##                               Dependent variable:
##                               -----
##                               limports
## -----
## lgdp_exporter                1.226***
##                               (0.008)
##
## lgdp_importer                0.951***
##                               (0.008)
##
## ldlist                       -1.374***
##                               (0.022)
##
## comlang_off                  1.293***
##                               (0.051)
##
## Constant                     -33.751***
##                               (0.350)
##
## -----
## Observations                 19,978
## R2                           0.642
## Adjusted R2                  0.642
## Residual Std. Error         2.492 (df = 19973)
## F Statistic                  8,969.064*** (df = 4; 19973)
## =====
## Note:                        *p<0.1; **p<0.05; ***p<0.01
```

## 参考文献

Arel-Bundock, Vincent. 2022. "Modelsummary: Data and Model Summaries in r." *Journal of Statistical Software* 103: 1–23. <https://doi.org/10.18637/jss.v103.i01>.

Blair, Graeme, Jasper Cooper, Alexander Coppock, Macartan Humphreys, Luke Sonnet, Neal Fultz, and Maintainer Graeme Blair. 2018. "Package 'Estimatr'." *Stat* 7 (1): 295–318. [https://search.r-project.org/CRAN/refmans/estimatr/html/lm\\_robust.html](https://search.r-project.org/CRAN/refmans/estimatr/html/lm_robust.html).