

R 入門 (4) 多重共線性、Panel 推定

1 多重共線性

```
setwd("/dail")
dat=read.table("04.csv",header = TRUE, sep =",")
attach(dat)
res1=cor(dat)
```

相関係数行列を res1 に保存

res1

	rent	space	distance	age	floor
rent	1.0000	0.8698	0.1976	-0.3650	0.1987
space	0.8698	1.0000	0.3683	-0.0682	0.1347
distance	0.1976	0.3683	1.0000	-0.0943	0.0279
age	-0.3650	-0.0682	-0.0943	1.0000	0.0393
floor	0.1987	0.1347	0.0279	0.0393	1.0000

```
install.packages("car")
library(car)
res2=lm(rent~space+distance+age+floor)
vif(res2)
```

Japan(Tokyo)を選択
library"car"を導入
重回帰を res2 に保存
res2 から VIF を計算

space	distance	age	floor
1.180163	1.163889	1.012578	1.021345

どれもほぼ 1

2 Panel 推定 (within 推定)

```
setwd("/dail")
install.packages("plm")
library(plm)
dat=read.table("14.csv",header = TRUE, sep =",")
attach(dat)
res1=lm(suicide ~ unemployment)
summary(res1)
```

OLS を res1 に保存

```
res2=plm(suicide~unemployment,data=dat,model="within",index=c("country_name","year"))
summary(res2)
```

oneway(州別効果のみ), index=c(id,時点)を指定

```
mu=fixef(res2)
summary(mu)
```

within を res2 に保存
固定効果を mu に保存

```
res3=plm(suicide ~ unemployment, data=dat, model="within", effect="twoways",index=c("country_name","year"))
summary(res3)
```

twoways(州効果と年効果), index=c(id,時点)を指定

```
pFtest(res2, res1) 制約なし res2 制約あり(州効果=0)res1 panel F 検定
```

F test for individual effects
data: suicide ~ unemployment
F = 207.67, df1 = 63, df2 = 1215, p-value < 2.2e-16

州効果有意

alternative hypothesis: significant effects

```
pFtest(res3, res2) 制約なし res3 制約あり(年効果=0)res2 panel F 検定
```

F test for twoways effects
data: suicide ~ unemployment
F = 9.4525, df1 = 19, df2 = 1196, p-value < 2.2e-16

時間効果有意

alternative hypothesis: significant effects

演習

テキスト p-246 の日本の自殺率と失業率の関係 (15.csv) について

(1) OLS, oneway, twoways の R の結果を示し、失業率→自殺率の大きさを比較しなさい

(2) Ftest をおこなって、OLS-oneway の比較で県効果、oneway-twoway の比較で年効果の有意性を検討し、採用すべきモデルを選びなさい

(1) OLS, oneway, twoways の比較

```

Call:
lm(formula = suicide ~ unemployment)

Residuals:
    Min      1Q  Median      3Q     Max 
-7.4742 -2.1887 -0.3819  1.7755 12.8032 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 12.0032   0.3902   30.76 <2e-16 ***
unemployment 2.4170   0.1108   21.81 <2e-16 ***  
---
Signif. codes:  0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.869 on 515 degrees of freedom
Multiple R-squared:  0.4802,    Adjusted R-squared:  0.4792 
F-statistic: 475.7 on 1 and 515 DF,  p-value: < 2.2e-16

One-way (individual) effect Within Model
Call:
plm(formula = suicide ~ unemployment, data = dat, model = "within",
index = c("prefecture", "year"))

Balanced Panel: n = 47, T = 11, N = 517

Residuals:
    Min. 1st Qu. Median 3rd Qu. Max. 
-4.947683 -0.958917 -0.031428  0.939817  5.860791 

Coefficients:
            Estimate Std. Error t-value Pr(>|t|)    
unemployment 3.06601   0.06957  44.071 < 2.2e-16 ***  
---
Signif. codes:  0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 5779.5
Residual Sum of Squares: 1124.2
R-Squared: 0.80549
Adj. R-Squared: 0.786
F-statistic: 1942.22 on 1 and 469 DF, p-value: < 2.22e-16

Two-way effects Within Model
Call:
plm(formula = suicide ~ unemployment, data = dat, effect = "twoways",
model = "within", index = c("prefecture", "year"))

Balanced Panel: n = 47, T = 11, N = 517

Residuals:
    Min. 1st Qu. Median 3rd Qu. Max. 
-4.477671 -0.718392 -0.019411  0.720720  5.551692 

Coefficients:
            Estimate Std. Error t-value Pr(>|t|)    
unemployment 0.87078   0.24055   3.62 0.0003274 ***  
---
Signif. codes:  0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 922.8
Residual Sum of Squares: 897.19
R-Squared: 0.027757
Adj. R-Squared: -0.092979
F-statistic: 13.1044 on 1 and 459 DF, p-value: 0.00032739

```

(2)

```

F test for individual effects

data: suicide ~ unemployment
F = 28.263, df1 = 46, df2 = 469, p-value < 2.2e-16
alternative hypothesis: significant effects

F test for two-way effects

data: suicide ~ unemployment
F = 11.612, df1 = 10, df2 = 459, p-value < 2.2e-16
alternative hypothesis: significant effects

```

失業率が自殺に及ぼす
プラスの影響は、どれも
有意だが、係数は 2.417
 $\rightarrow 3.066 \rightarrow 0.871$ のよう
に変化している。

こうした変化は欠落変
数バイアスによって生
じている。

結果だけで考察が書か
れてない or 考察だけで
結果が示されていない
レポートは低評価となりま
す

OLS<oneway

oneway<two-way

なので two-way モデル
が選択されるべき。効
果が大きいから
oneway を選ぶのは不
適切です