

第3章 重回帰分析

【課題】教科書 108-109 ページの例 5 のデータ（表 3-3）について、

- ① R を用いて重回帰分析をおこなってみる。
 - ② 相関係数行列と偏相関係数行列を求める。
- という 2 つのことをおこなってみよう。

1. Excelデータの作成

分析の準備として、Excel のデータファイルを作成しておく。ここでは、**ex3-3.csv** という名前で、**CSV 形式**で保存する。

	A	B	C	D
1	YEAR	CE	YD	YD1
2	1983	185234	219047	214658.4
3	1984	189610.5	223793.4	219047
4	1985	197124.2	230667.3	223793.4
5	1986	203577.3	236083.4	230667.3
6	1987	211994.4	240775.4	236083.4
7	1988	222311.8	253954.6	240775.4
8	1989	233017.5	265997.6	253954.6
9	1990	243711.2	279366.1	265997.6
10	1991	250592.4	291102.9	279366.1
11	1992	256900.5	295572.9	291102.9
12	1993	260282.1	297769.1	295572.9
13	1994	267276	302224.7	297769.1
14	1995	271955.8	304958.7	302224.7
15	1996	278575.1	305548.7	304958.7
16	1997	281151.7	309193	305548.7
17	1998	280228.3	311727.5	309193
18	1999	280262.4	311189	311727.5
19	2000	283412.8	310564.8	311189
20	2001	288066.6	306376.4	310564.8
21	2002	290377.5	309039	306376.4

2. Rプログラムの作成

「ファイル」－「保存」を選び、**ex3-3** という名前で保存する。ファイルの種類は **R files (*.R)** を選ぶ。プログラムの実行は「編集」－「すべて実行」である。そこに次のようにプログラムを入力していく。

```
data1 <- read.table("ex3-3.csv", header=TRUE, sep=",")
data1
reg1<-lm(CE~YD+YD1, data=data1)
summary(reg1)
windows()
plot(resid(reg1))
```

3. 自由度修正済み決定係数

重回帰分析の場合、説明変数の数を増やせば増やすほど、その説明変数が被説明変数と無関係であっても、決定係数の値は 1 に近づく。そのため、自由度修正済み決定係数で、モデルのあてはまりをチェックしなくてはならない。

「説明変数の数を増やせば増やすほど、その説明変数が被説明変数と無関係であっても、決定係数の値は 1 に近づく。」ことを、実際に確かめてみよう。**ex2-5.R** にスライドに示したような、阪神タイガース年間順位を加え、分析してみよう。

ex2-5.R の末尾に、次のような命令を加えればよい。

```
data1$ht <- c(4, 4, 1, 3, 6, 6, 5, 6, 6, 2, 4, 5, 6, 6, 5, 6, 6, 6, 4)
data1
reg2 <- lm(cr~yd+ht, data=data1)
summary(reg2)
windows()
plot(resid(reg2))
```

4. 相関係数、偏相関係数

R では 2 変量間の相関係数の行列を **cor** というコマンドで表示することができる。一方、偏相関係数は計算して導出する必要がある。

ex3-3.R を次のように修正し、**ex3-3c.R** という名前で保存してみよう。なお、#で始まる文はコメント行である。これらの行は R の操作に直接関係ないが、後で自分のプログラムを見て、何をやったのかを確認する時にあると便利である。

```
data1 <- read.table("ex3-3.csv", header=TRUE, sep=",")
data1
reg1<-lm(CE~YD+YD1, data=data1)
summary(reg1)
windows()
plot(resid(reg1))
#
#相関係数行列
#
cor(data1[,2:4])
#
#偏相関
#
par.cor<-function(x, y, z) {
  r.x.y<-cor(x, y)
  r.y.z<-cor(y, z)
  r.z.x<-cor(z, x)
  par.r<-(r.x.y-r.z.x*r.y.z)/((sqrt(1-(r.z.x)^2))*(sqrt(1-(r.y.z)^2)))
  return(par.r)
}
parcormat <- matrix(c(1, par.cor(data1$CE, data1$YD, data1$YD1), par.cor(data1$CE, data1$YD1, data1$YD), par.cor(data1$CE, data1$YD, data1$YD1), 1, par.cor(data1$YD, data1$YD1, data1$CE), par.cor(data1$CE, data1$YD1, data1$YD), par.cor(data1$YD, data1$YD1, data1$CE), 1), nrow=3, ncol=3)
parcormat
```

5. 分析結果の保存とWordへの読みこみ

R の分析結果はテキストファイルとして保存できるので、それをエディターやワープロソフトを使って読みこむことができる。また、グラフも画像形式で保存できるので、Word などで読みこむことが可能である。

ここでは、ex3-3c の出力結果と残差プロットを Word ファイルにまとめてみよう。

📖 手順

- ① 出力結果のウインドウをアクティブにして、「ファイル」－「ファイルを保存」を選び、**ex3-3c.txt** という名前で保存する。（ここで、デフォルトのファイル名は、lastsave.txt となっているが、適当に修正すること。）
- ② 残差プロットのウインドウをアクティブにして、「ファイル」－「別名で保存」を選ぶと、どの画像形式で保存するかを選択することになる。ここでは **bmp** を選び、**残差プロット** という名前で保存する。
- ③ Word を開き、**ex3-3c.txt** のファイルを読み出す。ファイルの種類を「すべてのファイル」にしておけば、すべての形式のファイルを読み出すことが可能である。
- ④ グラフファイルは、**挿入タブ**をクリックし、「図」ボタンをクリックして、**残差プロット**のファイルを選べば良い。グラフの大きさは、右クリックで「サイズ」を選べば変更できる。
- ⑤ 不要な部分を削除・修正し、タイトルと氏名、考察を加えればレポートが完成する。

☒ **演習問題** : ex3-3c の出力結果と残差プロットをまとめた Word ファイルに、考察を加えたレポートを完成させよう。