

Numeric Recipes for Econometrics(0)

入門

SHIMURA Masato

jcd02773@nifty.com

2010年12月2日

目次

1	3本のスクリプト-線形回帰入門	2
1.1	データとグラフ	2
1.2	3の簡潔なスクリプトの回帰モデル	8
1.3	多項式モデル	14
1.4	自己回帰モデル AR	16
1.5	関数一覧	19
2	計量経済の便利な小道具	20
2.1	平均と比較	20
2.2	成長率への応用	21
2.3	ラスパイレス指数	25
2.4	データの標準化	27
2.5	データの対数化	29
2.6	ノルム化	30
2.7	共分散行列と相関行列	31
2.8	相関係数と相関行列	34
2.9	偏相関行列	35
2.10	関数一覧	36
付録 A	行と列の取りだし	38
付録 B	EXCEL などスプレッド・シートのデータを用いる	39
B.1	入手	39

B.2	EXCEL 側でしておくこと	39
B.3	tara	39
B.4	CSV ファイル	41

概要

1 3本のスクリプト-線形回帰入門

単回帰、多項式回帰、自己回帰 (AR) の3のアルゴリズムの紹介である。何れも正規方程式は複雑な数式から導出されるが、正規方程式を経ないで直接データから、 $\frac{y}{X}$ となるような X を作成し、マトリクス X の逆行列を用いて行列除算を行うと、簡潔な一行のスクリプトで書き表すことができる。

1.1 データとグラフ

1.1.1 データの読み込み

本書で扱うデータはマトリクスを次のように $X_1 X_2 X_3 \dots X_n, Y$ の様に縦に並べる。(y が右端の行になる)

- 小さい Example データなどは次のようにスクリプトファイル (*ijs*) に書いておく
DN11

5.6 30
 – データを縦に並べるには *Stich(,.)* を最終行に用い 5.8 26
 る。他はカンマ (,) で連結する 6 33

6.2 31
 EX0=: 1 2 3 ,4 5 6 ,:7 8 9 6.4 33

- *ijs* ファイルに書き込むときは=:とグローバル定義を 6.4 35
 用いる。 6.4 37

6.6 36

6.8 33

*1

- 大きなデータは EXCEL からのインポートや CSV *2ファイルから出し入れする。

EXCEL や CSV ファイルの取扱いは APPENDIX に記載した。

(年度) (GDP) (民間消費) (民間投資) (政府支出)

94 473 272 65 131

95 484 278 67 138

*1 =は等しいかどうかの判別関数 (*equal*)

*2 Camma Separared Value カンマで一個ずつ区切ったファイル。中身は数値でなくテキストである。

96	502	286	74	142
97	504	283	79	137
98	500	285	75	134
99	503	285	75	137
0	515	287	81	141
1	509	289	78	137
2	514	291	76	139
3	524	293	82	137
4	534	296	86	137

*3

(出典) 伴・中村・跡田「エコノメトリクス」第2版 有斐閣 2006
次の例のように CSV ファイルを読み込む。

*4

```
require 'files csv plot'
] DN10=. ".@> readcsv '/data/excel/stat_j/csv/ban_1.csv'
```

(フォルダ/data/excel/stat_j/csv/に格納されているとき)

1.1.2 データの取り扱い

配列計算言語は概ね便利なインタプリッタでプログラムのコンパイルなどは一切不要であり、アナリストは自分の手で摘んで任意に解析できる。

数値と文字 数値と文字

EXCEL や CSV から読み込んだ状態ではボックスに入った文字であるので数値化してボックスから出さなければならない。

- ボックスのまま数値化 ⇒ ".(L:0) DN10
- マトリクスの構成を保ってボックスを開く ⇒ ;("1) (または ;("2))
- 合わせ技 ⇒ ".@>

縦と横

- J は 1 行のデータは横に数値を並べる。
- 本稿では 2 行以上の場合にはマトリクスとして縦に並べて取り扱う (統一)

データの取り返し (J の関数) J の行又は列の取り出し, 落としの関数

0 オリジンである。列に働かせるにはランク ("1) を用いる。

*3 データのサンプルは numeric.recipe_data.ijs をロードしたのち手動で読み込むようににしてあるがフォルダの構成が異なるとエラーになるのでフォルダの位置を修正して書き直す。

*4 CSV Comma Separated Value カンマで区切った (テキスト) データ。EXCEL や DB から CSV 形式で出力する。

```

Take      {
Take first { . 0 "1 DN11
Take last  { : { "1 DN11
Drop first { . } "1 DN11
Drop last  { : } "1 DN11

```

データの取り直し (作成) J の落としの関数は前から (後ろから) n 個と指定することはできるが、任意の行や列を指定できないので 配列を簡単に取り出し並び替えるユーティリティを作成した。^{*5}

<i>function</i>	<i>name</i>	<i>Usage</i>
0 オリジン	<i>puc</i>	<i>pickup column</i>
	<i>rec</i>	<i>remove column</i>
	<i>puc_set</i>	<i>pickup with set</i>
1 オリジン	<i>puc1</i>	<i>pickup column</i>
	<i>rec1</i>	<i>remove column</i>
	<i>puc_set1</i>	<i>pickup with set</i>

```

puc=:pick_column=: [ {"1 ]
puc1=:pick_column=: (>:@[) {"1 ]
rec=:remove_column=: 4 : '(I.-.(i.{$ y) e. x) puc y'
NB. remove_raw
rer=: remove_raw=: 4 : '(I.-.(i.# y) e. x) { y'

```

乱数の Example	列の取りだし	列の落とし
] a=. ?. 7 4 \$ 28	2 3 puc a	1 2 rec a
6 3 7 8	7 8	6 8
2 23 0 25	0 25	2 25
20 26 18 14	18 14	20 14
13 6 3 0	3 0	13 0
2 4 26 6	26 6	2 6
19 16 8 16	8 16	19 16
8 11 5 27	5 27	8 27

^{*5} 何番目という順序数で、行や列を数えるとき、0, 1, 2, 3 と数える 0 オリジン派と 1, 2, 3.. と数える 1 オリジン派がいる。本稿は 0 オリジンで統一しているが、1 オリジン派の為の関数も幾らか用意した。

最少自乗法は $\frac{y}{1, X_0, X_1, X_2, \dots, X_n}$ と X と y の 2 項関数である。次の 2 の方法を用意している。
(何れも 1 の付加はスクリプトで行う)

- `y % 1, X` と両項で計算する
- `1, X, y` と y を最後列に付加した一つのマトリクスにして単項で計算する (後はスクリプトに任せる)

最小自乗回帰 (単回帰、重回帰) の最もエレガントなスクリプトは次である

```
reg_ols=: %. 1&.,@]
```

回帰分析は従属変数が y , 説明変数が X である。J は左引数が x , 右引数が y であり紛らわしいので注意を要する

type 1 両項タイプ

- `y is 0 列, x is 1 列`

```
(0 puc a) reg_ols 1 puc a
```

5.72402 0.336313
 $f = 5.72402 + 0.336313x$
- `y is 0 列, x is 1 2 3 列 (重回帰)`

```
(0 puc a) reg_ols 0 rec a
```

8.92071 0.597347 _0.065022 _0.429712
 $y = 8.92071 + 0.597347x_0 - 0.065022x_1 - 0.429712x_2$

type 2 単項タイプ

- `y is 0 列 x is 3 列`

```
(0;3) puc_set a
```

	<code>x y</code>
<code>3 0 puc a</code>	8 6
<code>x y</code>	25 2
8 6	14 20
25 2	0 13
14 20	6 2
0 13	16 19
6 2	27 8
16 19	
27 8	

```
regx (0;3) puc_set a
```

11.1634 _0.0848279
 $y = 11.1634 - 0.0848279x$
- `y is 0 列 x is 1 3 列`

```
(0; 1 3) puc_set a
```

`x0 x1 y`

```

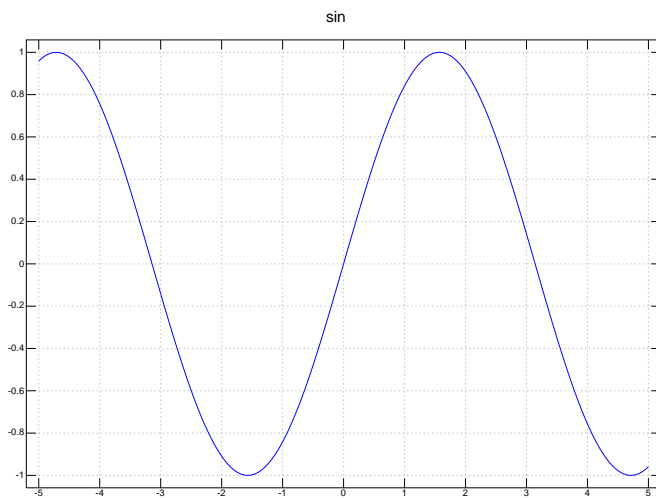
3 8 6
23 25 2
26 14 20
6 0 13
4 6 2
16 16 19
11 27 8
regx (0;1 3) puc_set a
8.05491 0.588484 -0.403745
 $f = 8.05491 + 0.588484x_0 - 0.403745x_1$ 

```

1.1.3 グラフ

*6

散布図やグラフを簡単に描くことができ、EPS、EMF などの形式で簡単に出力できる。
次の例はJの強力なサイエンスプロットの機能を用いている。



```

'title sin' plot _5 5 ;'1&o.' NB. sin from _5 to 5
pd 'eps /temp/sin0.eps' NB. save by eps

```

_5 5 plot の区間
(/temp に出力される)

J Grammar

*6 Jのグラフィックスは世界的に有名な会計システムに組み込まれている

数値化 ". format · 数値化 文字化は": ".@>はイディオム

ボックスを開く >open Box を開く <は Box 化

require 指定したシステムファイルをロードする

readcsv csv ファイルを読み込む 冒頭で require 'files csv' と指定する

plot グラフを描く

pd plot driver グラフの汎用コマンド

eps 画像形式の種類 ポストスクリプト形式

1.2 3の簡潔なスクリプトの回帰モデル

1.2.1 Matrix divide

重相関回帰、多項式回帰、自己相関 (AR) モデルを計算するスクリプトは次の3本の一行のスクリプトで表すことができる。

何れの式も中程に % . という記号が入っている。matrix - Divide という K.E.Iverson が創った強力な配列除算機能である。

*7

単回帰, 重回帰 Ordinary Least Square OLS	<code>regx=: 3 : '({:"1 y) % . 1,.: "1 y '</code>
多項式回帰 Polynomial	<code>poly1=:4 : 'y % . (>:i. # y)^/i. >: x '</code>
自己回帰 Auto Regression AR	<code>ar0=:4 : '(x }.tmp) % . .("1)}: >x <\ tmp=:y -(+/%#)y '</code>

1.2.2 OLS 回帰モデル

(単回帰)

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_n \end{bmatrix}, \begin{bmatrix} 1 & x_{11} \\ 1 & x_{12} \\ 1 & x_{13} \\ \vdots & \vdots \\ 1 & x_{1n} \end{bmatrix}$$

(重回帰)

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_n \end{bmatrix}, \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ 1 & x_{13} & x_{23} & \dots & x_{k3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix}$$

単回帰の正規方程式と重回帰モデルの正規方程式は次のように表すことができる。

$$(\mathbf{X}'\mathbf{X})\hat{\beta} = \mathbf{X}'\mathbf{y}$$

*7 APL ではドミノ母であらわす。最近の配列計算言語では / . と記述することが多い

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$$

これらは、更に次のように簡約される。

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y = \frac{\mathbf{X}'y}{\mathbf{X}'\mathbf{X}} = \frac{y}{\mathbf{X}}$$

線型数学で表現すれば単回帰も重回帰も同じである。配列計算言語では、単回帰も重回帰も同じスクリプトで記述でき、入力データにより、単回帰、重回帰のどちらにも適用できる。

定数項を求める場合には

$$\frac{y}{1, \mathbf{X}}$$

と \mathbf{X} の左に 1 を付加する。

Working Example 果実の直径と水分含有率

DN11

X ある果実の直径 *cm*

Y 水分含有率 (%) のデータ。

(データの出典 金谷 「これならわかる応用数学教室」)

サンプルデータ DN11 の逆行列を手計算で行おうとすれば大変である。教科書には正方行列の方法は紹介されているが、このような縦長の行列は先ず扱っていない。ここでは、素直に *K.E.Iverson* の *matrix - divide (%)* に従おう。

1.2.3 経過と説明

データは X,.Y の形に並べられている。	X に定数項の素の 1 を付加する	$\frac{y}{X} \rightarrow \frac{Y}{1,.X}$ 両項の %. はマトリクスの除算を行う
DN11	1,.0 puc DN11	
X Y	1 5.6	y %. 1,.X
-----	1 5.8	(1 puc DN11) %. 1,. 0 puc DN11
5.6 30	1 6	
5.8 26	1 6.2	_5.01128 6.03383
6 33	1 6.4	
6.2 31	1 6.4	reg0 DN11
6.4 33	1 6.4	_5.01128 6.03383
6.4 35	1 6.6	
6.4 37	1 6.8	y = -5.01128 + 6.03383x
6.6 36		
6.8 33		

*8

Script reg0

```

reg0=:3 : 0
NB. select trend data or multi data
if. 1= +/ * $ y do. reg_t y
elseif. do. regx y end.
)

reg_t=:3 : ' y %. 1,. >: i. # y '
regx=:3 : '({:"1 y ) %. 1,.}: "1 y '

```

1.2.4 定数項を用いない回帰

*8 データが X のみで Y は単に順序数でたりる場合には、reg_t を用いるが、reg0 は DN11 のような揃ったデータに用いる regx と自動で振り分けを行う。

定数項を用いない回帰

定数項を用いない場合は X の左に 1 のベクトルを付加しない。この場合は原点を通る。

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \frac{\mathbf{X}'\mathbf{y}}{\mathbf{X}'\mathbf{X}} = \frac{y}{X}$$

%. のみ用いる

```
({"1 DN11) %. {"1 DN11
```

```
5.23401
```

```
(1 puc DN11) %. 0 puc DN11
```

```
5.23401
```

```
y = 5.23401x
```

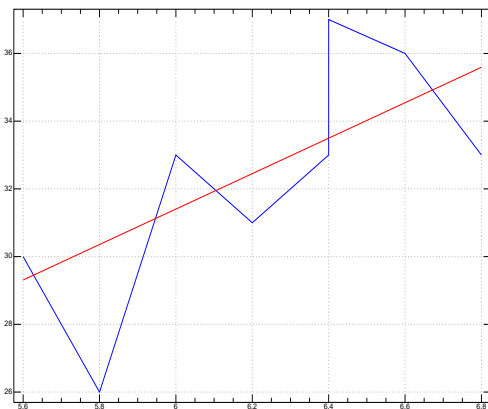


図 1 果実の直径と水分含有率

1.2.5 テスト

相関係数とモデルのテストは `reg_exam_ad` を `reg0` の後ろに連結して行う。

`reg_exam_ad` は副詞で動詞 (回帰関数) の強化のための検査車両の連結である

*9

```
reg0 reg_exam_ad DN11
```

```
+-----+-----+
```

```
|f=      |_5.01128 6.03383|
```

*9 動詞 (関数) を引数とする関数を J では副詞と言う

```

+-----+-----+
|corr=:|47.8237      | 相関係数
+-----+-----+
|AIC:  |18.8684      | AIC  情報量規準
+-----+-----+
|DW=   |2.17534       | ダービン/ワトソンの値
+-----+-----+
|t=:   |_0.33633 2.53299| t 値
+-----+-----+

```

1.2.6 ラインフィットと散布図

散布図とラインフィットは `linefit_reg0` で作成できる。

回帰係数を求めたなら原データと回帰した値をグラフに書いてみよう。

```

linefit_reg0 DN11
pd 'eps /temp/reg_0.eps' NB. Save

```

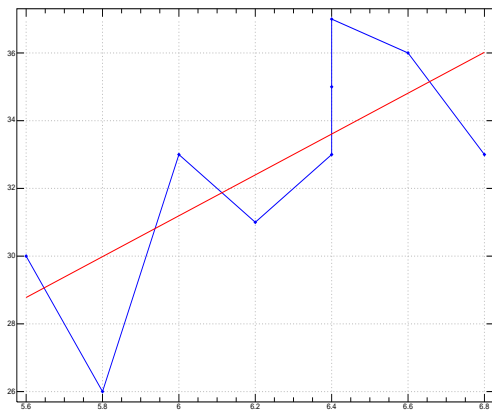


図2 果実の直径と水分含有率

1.2.7 J 言語ノート

副詞 関数 (数式を定義した動詞) を引数に取り専ら数式を計算する作用素 `1 : 0` で定義する

数列の生成 `i.` 数列を生成する `i.3` → 0 1 2 (3個) `i:3` → `_3 _2 _1 0 1 2 3`

`1` を足す `>: 1` を足す. `<: 1` を引く

行列除算 `%.` matrix divide 行列除算を行う 単項では逆行列を計算する

比較演算 `=` は左右の引数が等しいか否かを判定する。代入ではない

Σ +/は Σ +/ 1 2 3→ 6

1.3 多項式モデル

変数 y が t に関する k 次の多項式によって

$$y = C_0 + C_1 t + \dots + c_k t^k + \epsilon$$

のように表されている場合に、パラメータの最尤推定は次の式を解くことによって得られる。

$$\begin{bmatrix} S_0 & S_1 & \dots & S_k \\ S_1 & S_2 & \dots & S_{k+1} \\ S_2 & S_3 & \dots & S_{k+2} \\ \vdots & \vdots & \ddots & \vdots \\ S_K & S_{K+1} & \dots & S_{2k} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \dots \\ c_K \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ \dots \\ T_k \end{bmatrix}$$

多項式はフレキシブルで、次数を上手く選ぶとよくデータを追従する。逆に低次多項式はトレンドを概観でき、全体像を見るのに適している。^{*10}

Example ある年の2月の東京の最低気温（出典：鈴木義一郎）

$$y = 4.89554 + 1.13288x - 0.141456x^2 + 0.00438882x^3$$

3 poly0 DN12

4.89554 1.13288 _0.141456 0.00438882

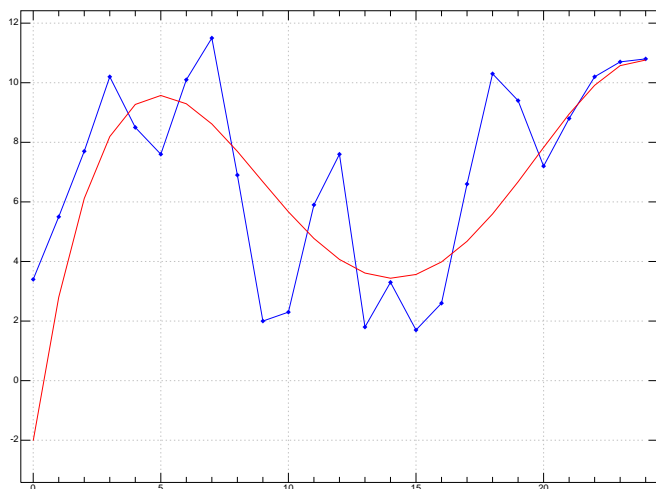


図3 東京の冬の最低気温 (青) と推計値 (赤)4 次

4 linefit_poly0 DN12

*10 次数により、両端が上下するという特徴があらわれる。

多項式の構成 データが長いと紙面を取るので、説明のため、10個の乱数データを用意した。

```
3 poly0 10?. 20
```

```
_11.2 17.6678 _3.75583 0.227855
```

$$f(x) = -11.2 + 17.6678x - 3.75583x^2 + 0.227855x^3$$

X, a = 10 ? . 20	X の 0 1 2 3 乗
<p>20 までの 10 個の乱数 X は単に順序数でオリジンを 1 とした。</p> <pre>(>:i. # a) , .a 1 6 2 3 3 19 4 15 5 10 6 14 7 0 8 7 9 12 10 17</pre>	<p>データの個数を求め、1 から順にナンバーを付け、これを 3 次の場合なら 0 1 2 3 乗する。0 乗が (1 の) 定数項になる。</p> <pre>(>:i. # a)^/ i. >: 3 1 1 1 1 1 2 4 8 1 3 9 27 1 4 16 64 1 5 25 125 1 6 36 216 1 7 49 343 1 8 64 512 1 9 81 729 1 10 100 1000</pre>

多項式の次数の選択 多項式の次数の選択は AIC を使い、AIC の値を最少にする次数 k を選択する。

AIC については第 3 章を参照

1.3.1 J 言語ノート

べき乗 \wedge

乱数 $?, ?$ はシードを固定して同じ乱数を生成する

データの個数 $\# n$ となる。 $\$$ はマトリクスの形を表す

1.4 自己回帰モデル AR

1.4.1 AR モデル

自己回帰モデル (AR :Auto Regressive Model) は離散型の時系列解析の代表的な方法であって、時系列の現時点における値が、自己の過去の時点の線形結合で表される。

M 次までの AR モデルの最少二乗解をすべて求めるには、次のようにすればよい。この行列を逆行列を用いればダイレクトに最少二乗解が得られる。

Autoregressive Model

$$\begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1k} \\ S_{12} & S_{22} & \cdots & S_{2k} \\ S_{13} & S_{23} & \cdots & S_{3k} \\ \vdots & \vdots & \ddots & \vdots \\ S_{1K} & S_{2K} & \cdots & S_{kK} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \cdots \\ b_K \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ \cdots \\ T_k \end{bmatrix}$$

自己回帰のアルゴリズムには *Yull – Walker* 法、*Burk* 法、*Householder* 法があるが、ここでは、データの逆行列から一度に計算してしまうシンプルな方法を紹介する

*11

前出の東京のある年の最低気温のデータ (DN12) のデータで 3 次の自己回帰の仕組みを再現してみよう。

*11 3 方式のアルゴリズムによる回帰係数には若干差があり、*Burk* 法はピークに強く、*Yull – Walker* 法は谷に強い。*householder* 法は中立的であると言われている。

1. $x_{t-1}, x_{t-2}, x_{t-3}, \dots, x_{t-n}$ という型のデータの組み合わせ (X) を作る。
2. 次数相当分のデータ (y) が組み合わせの結果減少するので、先頭から落として個数を合わせる。
3. $\frac{y}{X}$ を計算する。 X の逆行列が求まるので、簡単なマトリクスの除算で係数が求まる。

データ。	3次とする																																																																							
NR, y	$y, x_{t-0}, x_{t-1}, x_{t-2}, x_{t-3}$																																																																							
1 3.4	<pre>(3}.DN12),.}: . "1 >3<\DN12</pre> <table border="1"> <thead> <tr> <th>Y</th> <th>x_{t-1}</th> <th>x_{t-2}</th> <th>x_{t-3}</th> <th></th> </tr> </thead> <tbody> <tr> <td>3.4</td> <td>0</td> <td>0</td> <td>0</td> <td><i>cut</i></td> </tr> <tr> <td>5.5</td> <td>3.4</td> <td>0</td> <td>0</td> <td><i>cut</i></td> </tr> <tr> <td>7.7</td> <td>5.5</td> <td>3.4</td> <td>0</td> <td><i>cut</i></td> </tr> <tr> <td>10.2</td> <td>7.7</td> <td>5.5</td> <td>3.4</td> <td></td> </tr> <tr> <td>8.5</td> <td>10.2</td> <td>7.7</td> <td>5.5</td> <td></td> </tr> <tr> <td>7.6</td> <td>8.5</td> <td>10.2</td> <td>7.7</td> <td></td> </tr> <tr> <td>10.1</td> <td>10.1</td> <td>7.6</td> <td>8.5</td> <td></td> </tr> <tr> <td>11.5</td> <td>11.5</td> <td>10.1</td> <td>7.6</td> <td></td> </tr> <tr> <td>6.9</td> <td>6.9</td> <td>11.5</td> <td>10.1</td> <td></td> </tr> <tr> <td>2</td> <td>2</td> <td>6.9</td> <td>11.5</td> <td></td> </tr> <tr> <td>...</td> <td>...</td> <td>...</td> <td>...</td> <td></td> </tr> <tr> <td>10.7</td> <td>10.7</td> <td>10.2</td> <td>8.8</td> <td></td> </tr> <tr> <td>10.8</td> <td>10.8</td> <td>10.7</td> <td>10.2</td> <td><i>cut</i></td> </tr> </tbody> </table>	Y	x_{t-1}	x_{t-2}	x_{t-3}		3.4	0	0	0	<i>cut</i>	5.5	3.4	0	0	<i>cut</i>	7.7	5.5	3.4	0	<i>cut</i>	10.2	7.7	5.5	3.4		8.5	10.2	7.7	5.5		7.6	8.5	10.2	7.7		10.1	10.1	7.6	8.5		11.5	11.5	10.1	7.6		6.9	6.9	11.5	10.1		2	2	6.9	11.5			10.7	10.7	10.2	8.8		10.8	10.8	10.7	10.2	<i>cut</i>	<p><i>inifix()</i> の強力な機能で X を作成する。一つずつづらした組み合わせを <i>Box(<)</i> と <i>inifix()</i> で作り <i>Open(>)</i> で開く。</p> <p>$x_{t-1}, x_{t-2}, x_{t-3}$ の並びになるように縦に回転 <i>Rotate(."1)</i> させる。Y の時点と個数を合わせるため時点の古い上の方から指定次数分(3次)を落とす。</p> <p>最後の1行は回帰には用いないが、(虎の子として)予測に用いる。</p> <pre> . "1>3<\ >:i.5</pre> <pre>3 2 1</pre> <pre>4 3 2</pre> <pre>5 4 3</pre> <p>この形になれば重相関と同じようにマトリクスの除算(%.)で解が得られる。</p>
Y		x_{t-1}	x_{t-2}	x_{t-3}																																																																				
3.4		0	0	0	<i>cut</i>																																																																			
5.5		3.4	0	0	<i>cut</i>																																																																			
7.7		5.5	3.4	0	<i>cut</i>																																																																			
10.2		7.7	5.5	3.4																																																																				
8.5		10.2	7.7	5.5																																																																				
7.6		8.5	10.2	7.7																																																																				
10.1		10.1	7.6	8.5																																																																				
11.5		11.5	10.1	7.6																																																																				
6.9		6.9	11.5	10.1																																																																				
2		2	6.9	11.5																																																																				
...																																																																					
10.7	10.7	10.2	8.8																																																																					
10.8	10.8	10.7	10.2	<i>cut</i>																																																																				

1.4.2 AIC による次数選択

$$MLL = -\frac{n-k}{n}(\log Q_k/(n-k))$$

$$AIC = -2 \times MLL + 2 \times k = (n-k) \times (\log Q_k/(n-k)) + 2 \times k$$

AR の次数選択は AIC の値を最少にする次数 k を見つければよい。

```
4 exam_ar0 DN12
+-----+-----+-----+
|mean=6.904|corr=0.546834|AIC=41.4362|
+-----+-----+-----+

3 exam_ar0 DN12
+-----+-----+-----+
|mean=6.904|corr=0.536786|AIC=41.5126|
+-----+-----+-----+

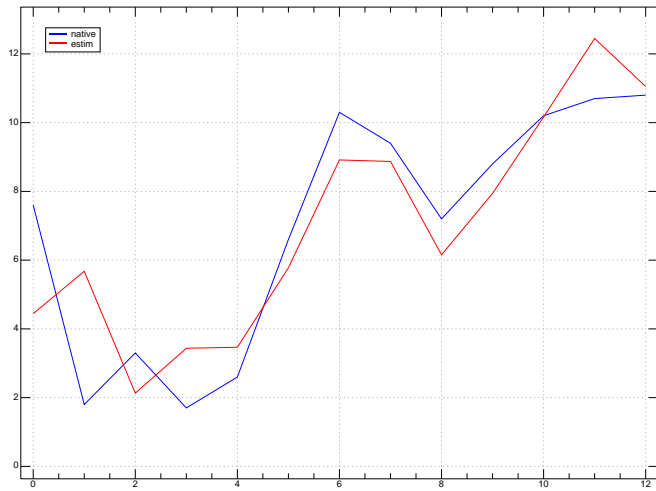
5 exam_ar0 DN12
+-----+-----+-----+
|mean=6.904|corr=0.56474|AIC=41.7886|
+-----+-----+-----+
```

```
4 ar0 DN12
0.973496 _0.548562 0.232687 0.0641936
```

4 次を採用する。

$$0.973496t_1 - 0.548562t_2 + 0.232687t_3 + 0.0641936t_4$$

```
4 linefit_ar DN12
pd 'eps \temp\ar_0.eps'
```



1.5 関数一覧

単回帰・重回帰	<i>reg0</i> <i>reg_exam_ad</i> <i>estim_reg0</i> <i>linefit_reg0</i>	<i>reg0 reg_exam_ad n</i> <i>estim_reg0 n</i> <i>linefit_reg0 n</i>
多項式回帰	<i>poly0</i> <i>poly_exam</i> <i>linefit_poly0</i>	<i>4 poly_exam n</i> <i>4 linefit_poly0 n</i>
自己回帰	<i>ar0</i> <i>exam_ar0</i>	<i>m ar0 n</i> <i>m exam_ar0 n</i>

2 計量経済の便利な小道具

データの前処理や比較、加工などに用いる小物ツールを集めた。各種の平均や成長率の求め方、分散共分散行列、相関行列やラスパイレス指数、グラフ表示に必携の標準化や対数化等々である。

2.1 平均と比較

2.1.1 平均

平均には日常用いる算術平均の他にも様々な方法がある。

```
]a=: >: i.10
1 2 3 4 5 6 7 8 9 10
```

項目	数式	Script	Example
算術平均	$\frac{1}{n} \sum_{i=1}^n X_i$	am=: +/ % # am2=: # %~ +/	am a 5.5
幾何平均	$\sqrt[n]{\prod_{i=1}^n X_i}$	gm=: # %: */	gm a 4.52873 データに0があると */ が0になり、マイナスがあると答えが複素数になるので使えない。
調和平均	$\frac{n}{\sum_{i=1}^n \frac{1}{X_i}}$	hm=: am &. (%"_)	逆数の平均値の逆数 比率 の平均に用いる hm a 3.41417

common mean 一般化平均	$\sqrt[m]{\frac{x_1^m + \dots + x_n^m}{n}}$	cm=: [: { . (am, gm) ^ : _	算術平均、幾何平均の極限 cm a 5.00257 更に一般化できる $f^{-1}\left(\frac{1}{n} \sum_{i=1}^n f x_i\right)$
----------------------	---	-----------------------------	--

2.1.2 調和平均-平均速度を求める

360 km の区間の最初の 120 km を時速30 km で、次の120 km を時速 40 km で、最後の120 km を60 km で走った場 合の平均速度を求める	hm 30 40 60 40 3 % +/- 1r30 1r40 1r60 40	$\frac{\frac{3}{1}}{\frac{4+3+2}{120}} = \frac{360}{9}$
--	---	---

2.2 成長率への応用

2.2.1 四半期データの前期比を年率換算

4 半期データの年率換算	$\left(\left(\frac{y_t}{y_{t-1}}\right)^4 - 1\right) \times 100$
--------------	--

2.2.2 四半期伸び率の年率換算

```
qtr_grow 1 1.02      NB. 1 1.02
8.24322
```

4 半期で 1.0 から 1.02 への伸びは年率換算 8.2%

2.2.3 Script

```
qtr_grow=: 3 : '100 * <: ^&4 %/|. y'
```

NB. e.g. u 1 1.02

NB. exchange rate of growth from quarter to year

経過と説明

新旧入れ替え 新旧を入れ替えて (rotate |.) から割る

```
%| . 1 1.02  
1.02
```

4 乗する ^4

1 を引く (<: は 1 を引く)

```
<: ^4 %| . 1 1.02  
0.0824322
```

100 倍する (* times)

```
100 * <: ^4 %| . 1 1.02  
8.24322
```

2.2.4 年平均成長率を求める。

5 年、10 年などの単位で、平均成長率を求める。

単位は % になっている。

年平均成長率	$\sqrt[n]{\frac{x_{t+n}}{x_t}} - 1 \times 100$
--------	--

Script

```
grow_ave=: 4 : '100 * <: (x %: %/ |. y)'
```

経過と説明

- 年次を逆にして比を求める。

```
%/ |. 504827 539160 NB. GDP 1995/2000  
1.06801
```

- 5 乗根を求め

```
5 %: %/ |. 504827 539160 NB. GDP 1995/2000  
1.01325
```

- 100倍する

100* 5 %: %/ |. 504827 539160 NB. GDP 1995/2000
101.325

Working Example

5 grow_ave 504827 539160 NB. GDP 1995/2000
1.32463 (1.32%)

10 grow_ave 469567 539160 NB. GDP 1990/2000
1.39161 (1.39%)

2.2.5 幾何平均-成長率の平均

成長率のデータしか手許にない場合に、成長率の平均が求められる。この場合幾何平均を用いる

DN20

1.25 1.4 1.07143 1.06667 1.125

gm DN20

1.17608

Working Example トヨタ自動車の年平均成長率

<i>year</i>	01	02	03	04	05
%	4.2	12.5	6.3	7.3	13.4

DN21

1.042 1.125 1.063 1.073 1.134

gm DN21

1.08681 NB. 8.67% per year

J言語ノート

<i>i.</i>		0 からの順序数を指定個数打ち出す
>:		1 を加える
*	×	かけ算
%	÷	割り算
*:	<i>square</i>	2 乗
%:	<i>square – root</i>	平方根両項は n 乗根
{.	<i>From</i>	最初のスカラ 行 列を取り出す
–	<i>infinity</i>	無限大
~	<i>child</i>	計算順序を逆にする
"1	<i>Rank</i>	マトリクスの演算方向 ランク 1 は行 (ベクトル) 方向
^:	<i>power</i>	一種のループ
&.	<i>under</i>	接続詞
y		右引数
.	<i>Rotate</i>	ベクトルやマトリクスを指定方向に回転させる
/	<i>Insert</i>	計算記号 (演算子) を数列の間に挿入する
<:	<i>Decriment</i>	1 を引く
&	接続詞	数字と演算子を繋いで動詞 (関数) を作る
r	分数	$1r3$ は $\frac{1}{3}$

2.3 ラスパイレス指数

指数は次により求められる。

$$2000 \text{ 年を } 100 \text{ とした指数} = \frac{y_t}{y_{00}}$$

幾つかの品目データを一つの指数にする方法に、ラスパイレス、パーシェ、フィッシャーの3の方法がありウエイトの置き方で区別される。ラスパイレス指数は自己のデータのウエイトで計算できるので便利である。

<i>Laspiress</i>	$\frac{\sum p_t q_0}{\sum p_0 q_0}$
<i>Parsh</i>	$\frac{\sum p_t q_t}{\sum p_0 q_t}$
<i>Fischer</i>	$\sqrt{\frac{\sum p_t q_0}{\sum p_0 q_0} \times \frac{\sum p_t q_t}{\sum p_0 q_t}}$

DN22

A 財 B 財
価格数量 価格数量

```
+-----+-----+
| 75 50|60 70| NB. 基準時
| 85 48|55 72| NB. 第1期
| 95 46|52 75| NB. 第2期
|105 44|48 80| NB. 第3期
+-----+-----+
```

データの入力は次のように行う。

```
DN22=(75 85 95 105,. 50 48 46 44) ;60 55 52 48,. 70 72 75 80
```

(データの出典 日本銀行調査統計局編 計量経済分析の基礎と応用 p 2 1)

```
lsp_chain0 DN22
```

```
1 1.01887 1.05535 1.08302 NB. 基準 第1期 第2期 第3期
```

```
par_chain0 DN22
```

1 1.01515 1.04025 1.04444

fis_chain0 DN22

1 1.01701 1.04777 1.06356

Working Example 生鮮野菜と1所帯あたり購入額

DN23=:13 170 15 155,52 49 55 41,13 89 14 85,51 57 53 53,41 55 35 57,:45 33 48 31

DN23=: ('';1 0 1 0)<;.1 DN23

DN23

2000 2005

価格数量価格数量

+-----+-----+

|13 170|15 155| NB. Cabbages

|52 49|55 41| NB. Spinachs

|13 89|14 85| NB. Napa

|51 57|53 53| NB. Leek

|41 55|35 57| NB. Lettuce

|45 33|48 31| NB. Broccoli

+-----+-----+

(出典)白砂第2版表2 - 14

lsp DN23

Las: Par: Fis:

103.654 103.209 103.431

2005年の野菜のラスパイレス、パーシェ、フィッシャー指数はそれぞれ上のとおり。

2.3.1 Script

lsp=: 3 : 0

NB. Calc Laspi Parshe Fischer

'P0 Q0 P1 Q1'=: {;("2),. |: L:0 y

las=.(+/ P1 * Q0)% +/ P0 * Q0

par=.(+/P1 * Q1)% +/ P0 * Q1

fis=.%:(las * par)

' Las: Par: Fis:',: 9j3 ":100 *(las, par ,fis)

)

Laspeyres *Étienne Laspeyres*(1834-1913)

Laspeyres はオックス語でオックス語を使うベルリンに移住したユグノーの家系の生まれ
経済と統計の教授としてベルン、リガ、カールスルーエを経て最後の 26 年はギーセンに落ち着
いた。有名な指数は 1871 年に発表した。(Wikipedia より)

```

n {      take      指定したスカラや行、列の取り出し
J 言語ノート  'Las'  character  文字列を入力 ( シングルクオート )
12j5 ":  format  文字化して出力する ( 12 桁内少数 5 桁)

```

2.4 データの標準化

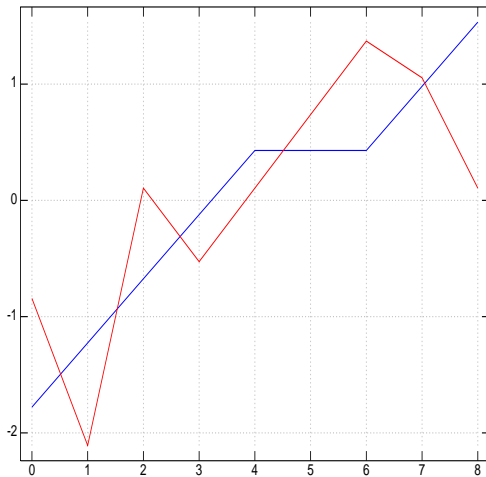
標準化とは、データを平均 0、分散 1 に変換することであり、単位が異なるデータの回帰や図示
に用いる。

標準化	$\frac{x - \bar{x}}{\sqrt{\frac{1}{n} \sum (x - \bar{x})^2}}$	データを平均 0、分散 1 に 変換する
-----	---	-------------------------

x ある果実の直径 cm		
y 水分含有率 (%) のデータ。	DN24	stand DN24
	5.6 30	_1.7781 _0.843274
	5.8 26	_1.22628 _2.10819
	6 33	_0.674453 0.105409
	6.2 31	_0.122628 _0.527046
	6.4 33	0.429198 0.105409
	6.4 35	0.429198 0.737865
	6.4 37	0.429198 1.37032
	6.6 36	0.981023 1.05409
	6.8 33	1.53285 0.105409

単位が異なるものはうまく一枚のグラフには表せないことが多いが標準化すると単位が無くなり、同一のグラフに表すことができる。

(データの出典 金谷 「これならわかる応用数学教室」)



```
'key fruits water'plot |: stand DN24
pd 'eps \temp\kanaya_02.eps'
```

図4 果実の直径(直線)と水分含有率

2.4.1 Script

```
stand=: dev % "1 sd
```

```
dev=: -"1(+/% #)
```

```
sd=: %:@var
```

```
var=: # %~ ([:+/[[: *: dev)
```

J 言語ノート

<code>*</code>	<i>square</i>	2 乗 (^2 と同じ)
<code>%</code>	<i>square_root</i>	2 乗根 ($\sqrt{\quad}$)
<code>-"1(+/%#)</code>	$x - \bar{x}$	残差
<code>[:</code>	<i>Cap</i>	関数を右から記述順に実行するように指定

2.4.2 平均を引く

平均を引く $x - \bar{x}$	-(+/%#)	時系列解析では前もってデータの平均を計算し、データから平均を引いた数を用いることがよく行われる。(最後に戻しておく)
------------------------	---------	--

2.5 データの対数化

重回帰では対数線形を用いることにより、不均一分散が取り除かれる、係数が弾性値を表し解釈しやすい、単位が取捨されるなどの好ましい性質があるので、頻繁に用いられる。底に e を用いる自然対数を用いる。log は \ln とも記述する。J には単項の \wedge で自然対数が、両項で、任意の数を底にとる対数のプリミティブが用意されている。

```

    ^ .100
^ 4.60517

^ 4.60517
100

```

自然対数 $e^y = x$ $\log x = y$	1x1 2.71828 J の 1x1 に e の数値が 組み込まれてある。	\wedge . DN24 1.72277 3.4012 1.75786 3.2581 1.79176 3.49651 1.82455 3.43399 1.8563 3.49651 1.8563 3.55535 1.8563 3.61092 1.88707 3.58352 1.91692 3.49651
-----------------------------------	---	---

$\hat{\cdot}$. *naturel log* 自然対数・常用対数は $10^{\hat{\cdot}}$. 100 is 2
 J 言語ノート $\hat{\cdot}$ e^n $\hat{\cdot}$ 4.60517 is 100
 $1x1$ e e の数値を組込んだ関数 $2x1$ $1x2$

2.6 ノルム化

$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$ はベクトルの大きさ (長さ) を表す。

長さの 2 乗の合計が 1 となるように基準化したユークリッド (Euclid) ノルムがよく用いられる。

ノルム化	<code>norm=: [: %: [: +/ *:</code>
------	------------------------------------

2.6.1 解説

The length (or norm) of \mathbf{v} is the nonnegative scalar $\|\mathbf{v}\|$

$$\|x\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

$$\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$$

$\|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$ は線分 PQ の長さをあらわす。

Example

$$\mathbf{v} = 1 \ -2 \ 2 \ 0 \quad \|\mathbf{v}\| = \sqrt{(1^2 - 2^2 + 2^2 + 0)} = \sqrt{9} = 3 = \frac{1}{\|\mathbf{v}\|} \mathbf{v}$$

```
euc_norm 1 _2 2 0
0.333333 _0.666667 0.666667 0
```

2.6.2 Script

```
norm=: [: %: [: +/ *:
euc_norm=: 3 : ' y. % norm y.'
```

Coffee Brake ネイピアと対数

対数とは底を等しくしたときの指数を取り出す関数である。スコットランド人でエディンバラの近くに生まれた *John Napier* (1550 – 1617) が「驚くべき対数法則の記述」を著したのは 1614 年であった。神学と星占術の研究者で数学者ではなかった。勃興してきた天文学の計算の工夫として、対数や e を考案し、天文学者を計算奴隷の仕事から解放した。またネイピアの計算棒といわれる計算器具も開発した。

2.7 共分散行列と相関行列

一度に分散共分散行列と相関行列を求めると、データ構造の見通しが良くなる。

2.7.1 分散と共分散

分散とはデータの平均値からの乖離を2乗したものの平均値である。

標準偏差とは、分散の正の平方根である。

残差平方和 (全変動)	$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i$ $\sum (x - \bar{x})^2$	ss=: [: +/ (*:@dev) mean=: +/ % # NB. am dev=: - mean
分散 Variance	$\frac{\sum (x - \bar{x})^2}{n}$	var=: ss%#
標準偏差 Standard deviation	$\sqrt{\frac{\sum (x - \bar{x})^2}{n}}$	sd=: %:&(ss%#)
変動係数	$\frac{sd}{\bar{x}}$	vr=: sd % mean
共分散	変数の相関の度合いを示す $\frac{1}{N} \sum (x - \bar{x})(y - \bar{y})$	cov=: # %~ ([: +/ [: */"1 dev) dev=: -"1(+/ % #) cov=: # %~ ([: +/ [:*/ "1 (-"1(+/ % #))) (一行で記述)

2.7.2 分散共分散行列と相関行列

多変数の場合、分散共分散行列や相関行列を一度に求めた方がデータの見通しがよい。

$$V(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^n V(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$$

$$V = \begin{pmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{pmatrix}$$

x	y
x1	y1
x2	y2
x3	y3
x4	y4

$$C = \frac{{}^tXX}{N} = \frac{1}{N} \begin{pmatrix} x_1 - \bar{x} & x_2 - \bar{x} & \cdots & x_n - \bar{x} \\ y_1 - \bar{y} & y_2 - \bar{y} & \cdots & y_n - \bar{y} \end{pmatrix} \begin{pmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ x_2 - \bar{x} & y_2 - \bar{y} \\ x_3 - \bar{x} & y_3 - \bar{y} \\ x_n - \bar{x} & y_n - \bar{y} \end{pmatrix}$$

$$\frac{1}{N} \begin{pmatrix} (x_n - \bar{x})(x_n - \bar{x}) & (x_n - \bar{x})(y_n - \bar{y}) \\ (y_n - \bar{y})(x_n - \bar{x}) & (y_n - \bar{y})(y_n - \bar{y}) \end{pmatrix} = \begin{pmatrix} x \text{ の分散} & \text{共分散} \\ \text{共分散} & y \text{ の分散} \end{pmatrix}$$

2.7.3 Worked Example

データ あるクラスの生徒	列の平均を引く
身長 体重 胸囲 座高	
DN25 NB. (1)	dev2 DN25 NB. (2)
145 30 60 70	_10 _10 _10 _10
145 35 70 75	_10 _5 0 _5
150 35 65 80	_5 _5 _5 0
150 40 70 70	_5 0 0 _10
155 40 75 75	0 0 5 _5
155 45 70 80	0 5 0 0
160 40 80 85	5 0 10 5
160 50 70 90	5 10 0 10
165 40 65 90	10 0 _5 10
165 45 75 85	10 5 5 5

<p>内積演算</p> <pre>(: dev2 DN25) +/ . * (dev2 DN25) NB. (3)</pre> <pre>500 275 175 425 275 300 150 250 175 150 300 100 425 250 100 500</pre>	<p>n で割る。n = 10</p> <pre>((: dev2 DN25) +/ . * (dev2 DN25)) % # DN25 NB. (4)</pre> <pre>50 27.5 17.5 42.5 27.5 30 15 25 17.5 15 30 10 42.5 25 10 50</pre>
<pre>variable DN25 50 27.5 17.5 42.5 27.5 30 15 25 17.5 15 30 10 42.5 25 10 50</pre>	

2.7.4 Script

```
variable=# %~ |: @dev2 +/ .* dev2
```

J 言語ノート +/ . * 内積
|: *Transpose* 転置・行と列の交換

2.8 相関係数と相関行列

先にデータを標準化しておくと同じ手順で相関行列が求められる。

$$R = \begin{pmatrix} 1 & r_{xy} & r_{xz} \\ r_{yx} & 1 & r_{yz} \\ r_{zx} & r_{zy} & 1 \end{pmatrix}$$

$$\rho_{xy} = \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$$

$$X, Y \text{ の相関係数 } \rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)} \cdot \sqrt{V(Y)}}$$

2.8.1 Working Example

```
cortable DN25
```

```

      1 0.710047 0.451848    0.85
0.710047      1    0.5 0.645497
0.451848    0.5      1 0.258199
    0.85 0.645497 0.258199      1

```

2.8.2 Script

```

cortable=: 3 : 0
ss=. [: +/ [: *: dev2
sd=. %:&(ss%#)
stand=. dev2@] %"1 sd@]
cortable=. #@] %~ ([:@stand@] +/ . * stand@])
cortable y
)

```

J 言語ノート @ *Atop* 動詞と動詞を繋いで一つの動詞とする
 - "1(+/%#) - "1 *is hook* データから縦の平均を縦方向に引く

2.9 偏相関行列

相関行列は、各変数間の相関係数の行列表示であり、データの特徴が一目で見られる。
 偏相関行列は、相関行列の逆行列とその対角行列から求められる。
 対角行列の外積をとり、その平方根で相関行列の逆行列 $-r_{ij}$ を割ればよい。

$$\text{相関行列 } R = [r_{ij}] \text{ の逆行列 } R^{-1} \text{ その要素 } r^{ij} r_{ij.o} = \frac{-r^{ij}}{\sqrt{r^{ii} r^{jj}}}$$

例えば X_1 と X_2 間の関係を残りの変数の影響を除去して計算したい場合には、偏相関を求める。
 この偏相関を、相関行列から一度に求めてしまう。

相関行列 $R = [r_{ij}]$
 R の逆行列 R^{-1} その要素 r^{ij}
 $r_{ij.o} = \frac{-r^{ij}}{\sqrt{r^{ii} r^{jj}}}$

相関行列	偏相関行列
<p>年間収入 5 分位の月当たりの所得と費用。</p> <p>DN26 単位円</p> <p>位 所得 教育 教養娯楽</p> <p>1 198005 2854 17662</p> <p>2 254020 5880 24208</p> <p>3 290237 11995 30538</p> <p>4 348839 19605 36568</p> <p>5 452356 23490 48113</p>	
<pre> cortable DN26 1 0.967571 0.997407 0.967571 1 0.976533 0.997407 0.976533 1 </pre>	<pre> pcor_table cortable DN26 1 _0.414878 0.965817 _0.414878 1 0.631025 0.965817 0.631025 1 </pre> <p>教育と教養娯楽の偏相関は 0.63</p>

2.10 関数一覧

標準化	<i>stand</i>	stand n
統計関数	<p>平均 <i>mean</i></p> <p>残差 <i>dev</i></p> <p>分散 <i>var</i></p> <p>標準偏差 <i>sd</i></p> <p>共分散 <i>cov</i></p>	

分散共分散行列 相関行列 偏相関行列	<i>vartable</i> <i>cortable</i> <i>pcor_table</i>	vartable n cortable n pcor_table cortable n
算術平均 幾何平均 調和平均	<i>am am2</i> <i>gm</i> <i>hm</i>	am i.10 gm i.10 hm i.10
四半期伸び率の年 率変換 年平均成長率	<i>qtr_grow</i> <i>grow_ave</i>	qtr_grow 1 1.02 5 grow_ave 504827 539160
ラスパイレス指数 ノルム	<i>lsp</i> <i>lsp_chain0</i> <i>par_chain0</i> <i>fis_chain0</i> <i>norm</i>	lsp n norm n

付録 A 行と列の取りだし

```
] a=. i. 5 4
0 1 2 3
4 5 6 7
8 9 10 11
12 13 14 15
16 17 18 19
```

- from (行) 0 1 3 _1 { a

```
0 1 3 _1 {a NB. rank is default
0 1 2 3
4 5 6 7
12 13 14 15
16 17 18 19
```

- from (列) 0 2 _1 {"1 a
ランクで縦方向 ("1) を明示する

```
0 2 _1 {"1 a
0 2 3
4 6 7
8 10 11
12 14 15
16 18 19
```

- take ({.)

```
2 {. "1 a
0 1
4 5
8 9
12 13
16 17
```

付録 B EXCEL などスプレッド・シートのデータを用いる

EXCEL や Libre CALC のファイルは biff-8 形式とする。(EXCEL2003 まで、EXCEL2007 は biff-12 である。biff は XML ではない)

EXCEL2003 までは動作確認 (2007 OK,2010?)

tara は Libre-Office のファイルも利用できる。(少し複雑)

B.1 入手

tara Net に繋がっている状態で J の Run/Package Manager を立ち上げ

tables/excel,tables/tara にチェックを入れて DL する

この環境にない場合は

- Net が利用できないときは J を CDROM にいれて,Net に繋がっているマシンを利用してセットアップする
- J602 のアドオンを利用している人から copy する

B.2 EXCEL 側でやっておくこと

- カンマを取り除く (数字の分離防止)
- データが空白のセルに 0 や-99999 などのダミーを入れておく (左詰め防止)

B.3 tara

最初に require 'files'

B.3.1 読み込み

tara.ijs tara.ijs を読み込む j602/addons/tables/tara/tara.ijs を読み込む

tutorial チュートリアル (tara.ijs) がある。

addons/tables/tara/tara.ijs

ファイルの読み込み .ロケールは指定しなくとも良い

dir =. '/data/sna/esri/principal/2010/' 適宜作成する

a=.readexcel dir,'shouhi_test.xls'

数値化 tara は既に数値化されている

Open BOX を開く

;("1) 9}. 2 4 {"1 a

Sheet を指定した読み込み .

```
'Sheet1' readexcel dir,'test_calc.xls'
+-----+
|1 |2 |3 | |
+-----+
|2 |3 |4 | |
+-----+
|4 |2 |6 | |
+-----+
|45|65|34| |
+-----+
```

B.3.2 書き込み

*12

オブジェクトの指定 最初に EXCEL のファイルの作成

```
bi=. '' conew 'biffbook'
```

書き込みの準備 準備後書き込む

```
writenumber__bi 0 0 ;i. 10 10
```

```
writenumber__bi 0 0 ;a1 NB. (example) a1=. ? 10 10 $ 100
```

```
save__bi '/temp/testtara.xls'
```

オブジェクトの指定は underbar を 2 個連ねる
(ファイルに) 書き込む こちらのほうが楽

```
a1=. i.4 5
```

```
a1 writexlsheets '/temp/tararest.xls'
```

```
1
```

*12 tara の書き込みと次の jmacros.xls での書き込みは J602 で非常にデリケートになっている。うまく行かない場合は csv が確実である

B.4 CSV ファイル

B.4.1 CSV ファイルを作成する

CSV Comma Separated Values

EXCEL 側でやっておくこと

- カンマを取り除く (数字の分離防止)
- データが空白のセルに 0 や-99999 などのダミーを入れておく (左詰め防止)
- copy して、数値のみのファイルにし、文字のエッジは捨てる
- コメントは別ファイルで

save EXCEL で csv を指定する

ファイル名は半角で

Example 消費総合指数を CSV 化する

消費総合指数は月次で 1994 からの時系列が毎月内閣府から公表される

<http://www5.cao.go.jp/keizai3/getsurei.html>

年や空白を全部埋めて上の index を全て取り除く

shouhi_test.csv で save する

B.4.2 CSV ファイルの利用

Jでの準備 require 'files csv'

読み込み dir=: 'c:/data/sna/esri/principal/2010/'

```
] a=. readcsv dir, 'shouhi_test.csv'
```

数値化] a=. ".@> readcsv dir, 'shouhi_test.csv'

ばらけるときは

```
] a=. ;("1) ".(L:0) a
```

References

金谷健一 これなら分かる応用数学教室 共立出版 2000

鈴木義一郎「J言語による統計計算」森北出版 1996

日本銀行調査統計局編 [計量経済分析の基礎と応用] 東洋経済新報社 1985

山澤成康 実践 計量経済学入門 日本評論社 2004

Miscellance

J602 is download available (No charge)

<http://www.jsoftware.com>

Scripts are accessible

<http://japla.sakura.ne.jp>