◆演習課題 6.E.1　第2章◆演習課題 2.E.1 のデータ(表 2.E.1)において，グループ変数 ID の影響を考量して，SA を独立変数，SB を従属変数として分析せよ．解答例は，著者のウェブサイトに挙げてある．

・http://y-okamoto-psy1949.la.coocan.jp/booksetc/pyda/

◆演習課題 6.E.2　第2章◆演習課題 2.E.2 のデータ(表 2.E.2)において，グループ変数 ID の影響を考量して，SA を独立変数，SB を従属変数として分析せよ．解答例は，著者のウェブサイトに挙げてある．

・http://y-okamoto-psy1949.la.coocan.jp/booksetc/pyda/

コラム 6.C.2　ダミー変数の独立性

演習課題 6.E.1 および 6.E.2 においてグループ変数をダミー変数として用意する方法は，第12章あるいは第15章のポアッソン回帰モデルの例を参考にすればよいが，第6章の重回帰モデルの場合，式(6.1.5)における逆行列が存在するために，X の列ベクトルが1次独立であるという前提がおかれている。

所属グループを表すダミー変数を XA，XB，XC と用意すると，以下に示すように1次独立ではなくなる。

X を

$$X=\begin{bmatrix} 1 & SA & XA & XB & XC \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & SA & XA & XB & XC \end{bmatrix}$$

と設定したとき，

$$\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}=\begin{bmatrix} XA \\ \vdots \\ XA \end{bmatrix}+\begin{bmatrix} XB \\ \vdots \\ XB \end{bmatrix}+\begin{bmatrix} XC \\ \vdots \\ XC \end{bmatrix}$$

が成り立ち，列ベクトルは独立ではない。

列ベクトルが独立であるようにダミー変数を設定するために，例えば，XA を除いて，XB と XC の2つを用いる。このとき，ダミー変数 XB と XC は，グループ A を基準にしたときの効果を表すために用いる。

第7章　主成分分析

7.1　モデル

データは複数の変量の値として与えられる．例えば，身長，体重，血圧の3つの変量の場合は，各人のデータは3次元ベクトル

(身長，体重，血圧)

で表される．このデータを40人から収集すると，40人のデータは3次元空間において40個の点として表される．変量の数が多い場合は変数の数の次元数の高次元空間の点として表されるが，これを低次元の空間に写してデータの分布が表されると，データの情報が読み取りやすくなる．例えば，2次元空間でデータの分布が表された場合は，平面上にデータの分布が散布図として表される．データの分布を低次元の空間で表す方法として主成分分析がある．

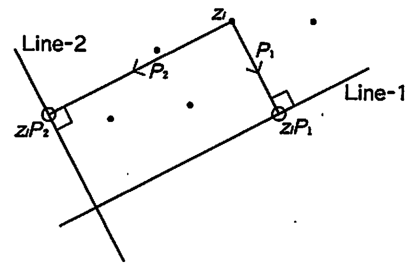

図 7.1.1　直線 Line-1 と Line-2 への2つの正射影

直感的に説明するために，図 7.1.1 のようにデータが2変量からなる場合を考える．詳しい説明は，Okamoto(2006)を参照されたい．このとき，データの各点は，平面上の分布として表される．図 7.1.1 では，5個のデータが表されている．これを1次元空間，すなわち直線上の分布として表すことを考える．図 7.1.1 では，直線 Line-1 と直線 Line-2 に点の射影をとる場合を表している．射影とは，点に光を当

ててその影を求めることと説明される．図 7.1.1 の場合は，直線に対して垂直真上から光を当てて直線上に影を求めているので正射影(orthogonal projection)と呼ぶ．影を落とす直線に対して斜め上から光を当てる場合は，射影(projection)と呼ぶ．

射影は，一般的には，s 次元空間から r 次元空間への影を求めることと考えられるが，正射影の場合は，r 次元空間に対して垂直な方向に光の方向が決まるので，影を投影する空間が決まれば，正射影も決まる．s 次元空間の点を

$$\boldsymbol{x}=(x_1 \ \cdots \ x_s) \tag{7.1.1}$$

と表したとき，r 次元空間への正射影は

$$\boldsymbol{x}P$$

と行列 P による積で表すことができる．主成分分析では，元の s 次元でのデータの分布を最もよく反映する r 次元への正射影を求める．元の分布を最もよく反映するということを，その分散が最大になることと考える．分散が最大になるということは，影の分布の広がりが最大になるということである．図 7.1.1 の例で言えば，直線 Line-1 への正射影が，2 次元での分布をよく反映するものと考える．直線 Line-1 に比べて直線 Line-2 への正射影は，正射影した点の散らばりが小さいという意味で，2 次元での分布の様子の反映が劣ると考える．

式(7.1.1)の形式のデータが n 組与えられているとする．それを行列

$$X=\begin{bmatrix}\boldsymbol{x}_1\\ \vdots \\ \boldsymbol{x}_n\end{bmatrix}=\begin{bmatrix}x_{11} & \cdots & x_{1s}\\ & \vdots & \\ x_{n1} & \cdots & x_{ns}\end{bmatrix} \tag{7.1.2}$$

で表す．主成分分析では変量間の関係を見るので，各変量を平均 0，標準偏差 1 に揃えて原点と単位の影響を除いて分析することが考えられる．式(7.1.2)の変量の値を平均 0，標準偏差 1 の標準得点に変換した式(7.1.3)について考える．

$$Z=\begin{bmatrix}\boldsymbol{z}_1\\ \vdots \\ \boldsymbol{z}_n\end{bmatrix}=\begin{bmatrix}z_{11} & \cdots & z_{1s}\\ & \vdots & \\ z_{n1} & \cdots & z_{ns}\end{bmatrix} \tag{7.1.3}$$

この $\boldsymbol{z}$ に対して，射影された ZP の分散が最大になる正射影 P を求める．r 次元空間での分散を次式(7.1.4)で表す．

$$tr\left\{\frac{1}{n}(ZP)'(ZP)\right\} \tag{7.1.4}$$

式(7.1.4)を最大にする正射影 P は，次式(7.1.5)で与えられる．

$$P=V_1V_1' \tag{7.1.5}$$

ここで，V_1 は次式の特異値分解で与えられるものである．

$$\frac{1}{\sqrt{n}}Z=U\Lambda V'=[U_1U_2]\begin{bmatrix}\Lambda_1 & 0\\ 0 & \Lambda_2\end{bmatrix}\begin{bmatrix}V_1'\\ V_2'\end{bmatrix}$$

$$U_1=[\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_r],\ U_2=[\boldsymbol{u}_{r+1} \ \cdots \ \boldsymbol{u}_s]$$

$$\Lambda_1=\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_r\end{bmatrix},\ \Lambda_2=\begin{bmatrix}\lambda_{r+1} & & 0\\ & \ddots & \\ 0 & & \lambda_s\end{bmatrix},\ \lambda_1\geq\lambda_2\geq\cdots\geq\lambda_s$$

$$V_1=[\boldsymbol{v}_1 \ \cdots \ \boldsymbol{v}_r],\ V_2=[\boldsymbol{v}_{r+1} \ \cdots \ \boldsymbol{v}_s]$$

このとき，

$$tr\left\{\frac{1}{n}(ZP)'(ZP)\right\}=\lambda_1^2+\cdots+\lambda_r^2 \tag{7.1.6}$$

である．式(7.1.6)の値は，$r=s$ のとき，

$$tr\left\{\frac{1}{n}(ZP)'(ZP)\right\}=\lambda_1^2+\cdots+\lambda_s^2=s \tag{7.1.7}$$

となる．

相関行列を R とおくと，

$$R=\frac{1}{n}Z'Z=V\Lambda^2V' \tag{7.1.8}$$

となり，相関行列の固有分解により Λ と V を得ることができる．

射影された ZP の空間において，座標軸を ZP の座標値の分散が最大になるように互いに直交するものを選ぶと $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$ になる．このときの座標値は主成分(principal components)と呼ばれている．主成分を C で表すと，次式で与えられる．

$$C=\sqrt{n}\,U_1=ZV_1\Lambda_1^{-1}$$

空間 ZP における基底を $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$ 以外にとったときは，ZP の座標値は成分(components)と呼ばれる．成分と ZP の関係を表す行列はパターン行列(pattern matrix)と呼ぶ．

主成分のパターン行列 P_{at} は次式で与えられる．

$$P_{at}=V_1\Lambda_1$$

空間 ZP における基底を $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$ 以外にとると，各座標軸の解釈が容易になる場合がある．いま，基底を $\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_r\}$ にとったとする．座標間の変換を表す行列を

T とおく．すなわち，

$$\begin{bmatrix} b_1' \\ \vdots \\ b_r' \end{bmatrix} = T \begin{bmatrix} v_1' \\ \vdots \\ v_r' \end{bmatrix}$$

である．この T は，回転(rotation)と呼ばれている．正規直交基底間の回転のときは，直交回転(orthogonal rotation)，それ以外のときは斜交回転(oblique rotation)と呼ばれる．

基底 $\{b_1, \cdots, b_r\}$ における成分 C_T とパターン行列 P_T は次式で与えられる．

$$C_T = C\Lambda_1 T^{-1}$$
$$P_T = P_{a1}\Lambda_1^{-1}T'$$

2つのパターン行列 P_{a1} と P_T が与えられると，回転 T は次式

$$T = P_T'P_T(P_{a1}'P_T)^{-1}\Lambda_1$$

で与えられる．

したがって，パターン行列 P_{a1} と P_T が与えられたときの成分の変換式は

$$C_T = C\Lambda_1 T^{-1} = CP_{a1}'P_T(P_T'P_T)^{-1}$$

となる．

主成分分析の具体例による説明を次節で行う．

7.2　Python スクリプト

本書用に用意した Python スクリプトは書籍中に掲載するには少し長すぎるので，一部の記載になるが，完全なスクリプトファイルは著者のウェブサイトからダウンロードできる．

・http://y-okamoto-psy1949.la.coocan.jp/booksetc/pyda/

主成分分析の基本的な部分について，表 7.2.1 のデータの分析を例に説明する．

表 7.2.1 のデータは，6 教科の 20 人分の得点である(仮想データ)．6 次元空間での 20 個の点として表せる．これを低次元の空間に正射影して教科間の得点の関係を調べる．

表 7.2.1　6 教科の得点(仮想データ)

通番 (ID)	国語 (Jpn)	英語 (Eng)	歴史 (Hist)	数学 (Math)	物理 (Phys)	化学 (Chem)
1	48	46	60	47	68	44
2	54	49	63	61	49	45
3	48	55	47	49	49	47
4	83	78	85	79	66	63
5	56	49	56	78	70	73
6	78	86	76	67	67	75
7	62	64	59	62	69	57
8	80	78	82	82	80	75
9	55	49	58	79	67	75
10	53	65	63	75	89	77
11	45	52	53	47	41	42
12	59	62	72	56	72	72
13	44	34	49	54	65	59
14	79	64	75	37	15	9
15	62	69	66	71	67	66
16	72	63	66	80	84	72
17	78	82	83	81	67	81
18	62	73	63	60	64	67
19	66	60	61	64	43	63
20	70	66	71	50	55	52

まず，データの読込みであるが，これは次のスクリプトが示すように，テキストファイルからの読込みと csv ファイルからの読込みの 2 通りを用意している．

```
print('Your data is in a Text File...Choose 1')
print('Your data is in a CSV File....Choose 2')
ck_choice = input('Your choice = ')
data = None
if ck_choice == '1':
    data, s = ReadTextFile()
elif ck_choice == '2':
    data, s = ReadCSVFile()
```

テキストファイルとして用意するデータは，図 7.2.1 に示す形式で作成する．図 7.2.1 のデータファイルは，表 7.2.1 のデータに対するものである．データは，スラッシュ '/' で始まる行で挟まれている．スラッシュ '/' で始まる最初の行に続いて，変量の数が書かれ，次の行に変量のラベルが並べられている．変量の先頭は，各データの識別用文字列(図 7.2.1 の場合は，通し番号を表す数値文字列)の名前であり，

その後，主成分分析の対象となる変量の名前が並べられている．変量の名前を並べた次の行から，1行に1ケースずつのデータを並べる．先頭は，ケースIDの文字列で，それに続いて各変量のデータを書き並べる．データの最後は，スラッシュ'/'で始まる行をおいて，データの終わりであることを示す．

DataSubjects.txt - M:¥Books¥PythonDA¥サンプル0420¥主成分分析¥主成分分析スクリプト¥pcafiles¥DataSubjects.txt (3.6.5)

File Edit Format Run Options Window Help

```
/
      6
     ID     Jpn     Eng     Hist     Math     Phys     Chem
      1      48      46      60       47       68       44
      2      54      49      63       61       49       45
      3      48      55      47       49       49       47
      4      83      78      85       79       66       63
      5      56      49      56       78       70       73
      6      78      86      76       67       67       75
      7      62      64      59       62       69       57
      8      80      78      82       82       80       75
      9      55      49      58       79       67       75
     10      53      65      63       75       89       77
     11      45      52      53       47       41       42
     12      59      62      72       56       72       72
     13      44      34      49       54       65       59
     14      79      64      75       37       15        9
     15      62      69      66       71       67       66
     16      72      63      66       80       84       72
     17      78      82      83       81       67       81
     18      62      73      63       60       64       67
     19      66      60      61       64       43       63
     20      70      66      71       50       55       52
/
```

図 7.2.1　テキストファイルデータの例

データをcsvファイルとして用意する場合は，まず，図7.2.2のように準備する．図7.2.2は，表7.2.1のデータに対するものである．1行目のセルに変量名を入れ，2行目からデータを入れていく．データの設定が終われば，「名前を付けて保存」メニュで，「ファイルの種類」を「CSV（コンマ区切り）(*.csv)」を選んで保存する．この場合，Excelではいろいろなことが確認されるが，すべてYesを選んで先に進めばよい．保存したファイルをテキストエディタで開くと，図7.2.3のように，セル内に設定した値がコンマで区切られて並んでいることがわかる．

リスト7.2.1にデータファイルの読込みモジュールを示す．

	A	B	C	D	E	F	G	H
1	ID	Jpn	Eng	Hist	Math	Phys	Chem	
2	1	48	46	60	47	68	44	
3	2	54	49	63	61	49	45	
4	3	48	55	47	49	49	47	
5	4	83	78	85	79	66	63	
6	5	56	49	56	78	70	73	
7	6	78	86	76	67	67	75	
8	7	62	64	59	62	69	57	
9	8	80	78	82	82	80	75	
10	9	55	49	58	79	67	75	
11	10	53	65	63	75	89	77	
12	11	45	52	53	47	41	42	
13	12	59	62	72	56	72	72	
14	13	44	34	49	54	65	59	
15	14	79	64	75	37	15	9	
16	15	62	69	66	71	67	66	
17	16	72	63	66	80	84	72	
18	17	78	82	83	81	67	81	
19	18	62	73	63	60	64	67	
20	19	66	60	61	64	43	63	
21	20	70	66	71	50	55	52	
22								
23								

図 7.2.2　csvファイルの準備(Excel)

DataSubjects.csv - M:¥Books¥PythonDA¥サンプル0420¥主成分分析¥主成分分析スクリプト¥pcafiles¥DataSubjects.csv (3.6.5)

File Edit Format Run Options Window Help

```
ID,Jpn,Eng,Hist,Math,Phys,Chem
1,48,46,60,47,68,44
2,54,49,63,61,49,45
3,48,55,47,49,49,47
4,83,78,85,79,66,63
5,56,49,56,78,70,73
6,78,86,76,67,67,75
7,62,64,59,62,69,57
8,80,78,82,82,80,75
9,55,49,58,79,67,75
10,53,65,63,75,89,77
11,45,52,53,47,41,42
12,59,62,72,56,72,72
13,44,34,49,54,65,59
14,79,64,75,37,15,9
15,62,69,66,71,67,66
16,72,63,66,80,84,72
17,78,82,83,81,67,81
18,62,73,63,60,64,67
19,66,60,61,64,43,63
20,70,66,71,50,55,52
```

図 7.2.3　csvファイルをテキストエディタで開いた場合

リスト 7.2.1　データ読込みモジュール readfile.py

```
import csv

def ReadTextFile():
    #
    #      Prepare the input data file
    #
    s = input("input data file (*.txt) = ")
    f = open(s, "r")
    #
    #      Set the contents of the input data file in the object data
    #
    data = f.readlines()
    f.close()
    return data, s

def ReadCSVFile():
    #
    #      Prepare the input data file
    #
    f_in_nm = input("input data file (*.csv) = ")
    #    f_in_nm = input('input data file (*.csv) = ')
    with open(f_in_nm, 'r') as f:
        csv_data = [d for d in csv.reader(f)]
    #
    #      Set the contents of the csv file in the data
    #
    data = [' ']
    data.append('/')
    data.append('    {}'.format(len(csv_data[0]) - 1))
    for i in range(len(csv_data)):
        temp_str = ''
        if i == 0:
            for v in csv_data[i]:
                temp_str += ' ' + v
        else:
            for j in range(len(csv_data[i])):
                if j == 0:
                    temp_str += ' ' + csv_data[i][0]
                else:
                    temp_str += '    {}'.format(csv_data[i][j])
        data.append(temp_str)
    data.append('/')

    return data, f_in_nm
```


変量の値を行列 X に設定した後，標準化得点を次のスクリプトで Z に求めている．

```
import scipy.stats as scst
Z0 = []
for j in range(n_vars):
    Z0.append(scst.zscore(X[j]))
Z = np.array(Z0).transpose()
```

求めた Z から相関行列 R を計算して，固有分解を行い，Λ^2 と V を求めることを，次のスクリプトで行っている(式(7.1.8)参照)．

```
R = (Z.transpose() @ Z) / n_data
Lmbd2, V = eigen_sym(R)
```

求めた固有値と固有値の 2 乗の累積和が図 7.2.4 のように表示される．式(7.1.7)より，固有値の 2 乗のすべての和は，変量の総数 s，いまの場合は 6 に等しい．したがって，有用な固有値は，固有値の 2 乗の総和を変量の総数で割った値である 1 より大きいものという基準を考えることもできる．図 7.2.4 の場合，1 より大きい固有値は 2 個である．成分の個数を 2 と入力して Enter キーを押す．指定した主成分数で計算が始まり，終了するとパターン行列の値を座標とする変量の散布図を描くときの主成分の選択が求められる(図 7.2.5)．横軸と縦軸の主成分を選ぶと，それに対応する散布図が表示される(図 7.2.6)．主成分において正負の方向は固有分解のときの計算法に依存して決まるもので，正負を逆転したものも正しい解である．すなわち，図 7.2.6 において，座標軸を正負反転してもよい．図 7.2.6 の散布図では，横軸がデータの主要な傾向を表す成分であることがわかる．一般に，主成分分析において，第 1 主成分は固有値の最大であるものに対応していて，データの最も主要な傾向を表す．図 7.2.6 では，科目得点の全体的傾向，すなわち，負の方向に絶対値が大きくなるほど 6 教科全体の成績がよいことを表している．縦軸は文系・理系科目の区別に対応している．

図 7.2.6 のフォームの右上の×印をクリックして閉じると，次の表示用成分の指定が求められる(図 7.2.7)．「xdim =」に 1 より小さい値を設定すると，主成分によるパターン行列の表示は終わり，次の回転後の成分に対する表示に移る．回転は，各成分が変数のグループの中心を通るように行われる．図 7.2.6 で言えば，理系の科目の集まりと文系の科目の集まりの中央に各成分を表す軸が通るように回転が行われる．回転は数学的に目的関数を設定して行われ，いろいろなものが提案されて

```
Your data is in a Text File...Choose 1
Your data is is a CSV File....Choose 2
Your choice = 1
Input data file (*.txt) = DataSubjects.txt
Output data file = Results.txt


                      Lambda   cum.sqr        %

            Lambda-1  1.83432  3.36473    56.08
            Lambda-2  1.40444  5.33718    88.95
            Lambda-3  0.53507  5.62349    93.72
            Lambda-4  0.46458  5.83932    97.32
            Lambda-5  0.31091  5.93599    98.93
            Lambda-6  0.25301  6.00000   100.00

Number of components =
```

図 7.2.4　固有値 λ_i の出力と求める主成分の数の設定

```
Number of components = 2

Set the component number (from 1 to 2) to plot the pattern.
If you do not want to plot the pattern, set the number less than 1.
xdim = 1
ydim = 2
```

図 7.2.5　表示主成分の選択

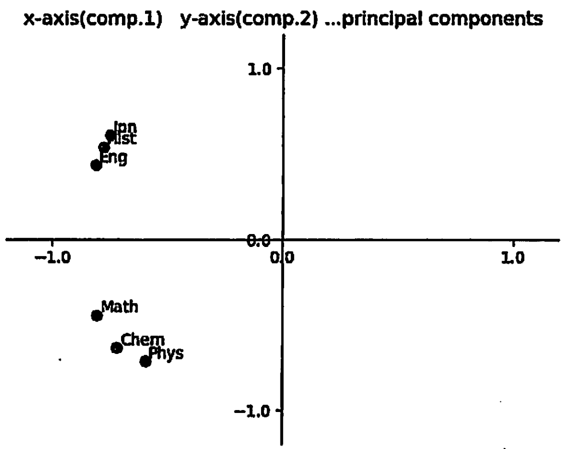

図 7.2.6　パターン行列の描画

いるが，本スクリプトでは varimax 回転と呼ばれている方法が使われている．図 7.2.7 では，varimax 回転を行ったときの第 1 成分と第 2 成分が指定されている．

このときのパターン行列の散布図を図 7.2.8 に示す．横軸の負の方向が理系科目の得点，縦軸の正の方向が文系科目の得点を表していることがわかる．

図 7.2.8 のフォームの右上の×印をクリックして閉じると，次の表示用の成分の指定が求められる．「xdim =」に 1 より小さい値を設定すると，varimax 回転後のパターン行列の描画表示は終わり，次の斜交回転が行われる．斜交回転の場合は，各軸が直交していないので，直交座標による表示は行われずにスクリプトの実行終了となる．

スクリプトの実行終了後，出力ファイルを開くとリスト 7.2.2 のようになっている．読み込んだデータとその標準化得点が出力されている．続いて，λ_i と λ_i^2 の累積和，

```
Number of components = 2

Set the component number (from 1 to 2) to plot the pattern.
If you do not want to plot the pattern, set the number less than 1.
xdim = 1
ydim = 2

Set the component number (from 1 to 2) to plot the pattern.
If you do not want to plot the pattern, set the number less than 1.
xdim = 0

Set the component number (from 1 to 2) to plot the pattern for the Varimax criterion.
If you do not want to plot the pattern, set the number less than 1.
xdim = 1
ydim = 2
```

図 7.2.7　次の varimax 回転後のパターン行列の表示における成分の選択

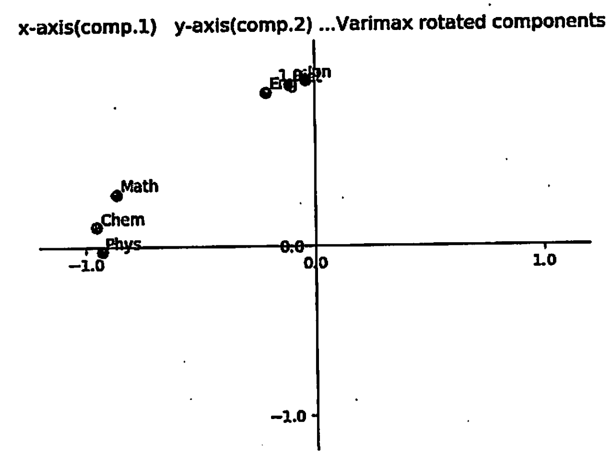

図 7.2.8　varimax 回転後のパターン行列の表示

および累積和の全体に対する比率が出力されている．第2固有値までが1より大きいので，データの主要な傾向を示すものであると考えられる．その後，主成分に対するパターン行列が出力されている．第1主成分に対する値がすべて負であるので，第1主成分の負の値が科目全体の得点傾向を表している．第2主成分については，文系が正，理系が負であるので，文系／理系の区別を表している．各ケースの主成分の値の出力の後，varimax 回転後のパターン行列が出力されている．第1成分については，文系科目の値が0に近く，理系科目の値は絶対値が1に近いので，第1成分の負の値は理系科目の得点に対応していると解釈できる．第2成分については，文系の科目の値が1に近く，理系科目は0に近いので，文系科目の得点を表す成分であると考えられる．パターン行列については，パターンが見やすくなるように変数を並べ替えたものも出力されている．表 7.2.1 のデータの場合は，並替えを行わなくても解釈が容易であるが，データによっては並べ替えることにより解釈が容易になることが多い．各ケースの成分の値の出力後，斜交回転を行った結果が出力されている．斜交回転は promax 法と呼ばれている回転が用いられている．斜交回転では軸の直交条件がないので，パターン行列がより単純化される．すなわち，直行回転に比べてパターン行列の絶対値がより1あるいは0に近いものになる．軸が直交していないので，各成分の相関が0ではなくなる．成分間の相関行列が出力されている．

リスト 7.2.2 計算結果の出力例

```
Input data file = DataSubjects.txt


n_vars = 6
Var_names...
    ID    Jpn     Eng     Hist    Math    Phys    Chem

     1  48.000  46.000  60.000  47.000  68.000  44.000
     2  54.000  49.000  63.000  61.000  49.000  45.000
     3  48.000  55.000  47.000  49.000  49.000  47.000
           .
           .
           .
    18  62.000  73.000  63.000  60.000  64.000  67.000
    19  66.000  60.000  61.000  64.000  43.000  63.000
    20  70.000  66.000  71.000  50.000  55.000  52.000
```



```
Standardized data...

  ID      Jpn       Eng      Hist      Math      Phys      Chem
   1  -1.20423  -1.24409  -0.50790  -1.26076   0.34673  -0.99764
   2  -0.71271  -1.01371  -0.22573  -0.21942  -0.81927  -0.93790
   3  -1.20423  -0.55293  -1.73062  -1.11200  -0.81927  -0.81842
           .
           .
           .
  18  -0.05734   0.82940  -0.22573  -0.29381   0.10126   0.37636
  19   0.27034  -0.16895  -0.41384   0.00372  -1.18748   0.13740
  20   0.59802   0.29182   0.52671  -1.03762  -0.45106  -0.51973

Lambda...
              Lambd    cum.sqr       %
  Lambda-1   1.83432   3.36473    56.08
  Lambda-2   1.40444   5.33718    88.95
  Lambda-3   0.53507   5.62349    93.72
  Lambda-4   0.46458   5.83932    97.32
  Lambda-5   0.31091   5.93599    98.93
  Lambda-6   0.25301   6.00000   100.00


Pattern matrix for the principal components =
          comp.1    comp.2
   Jpn  -0.75184   0.61221
   Eng  -0.81468   0.43846
  Hist  -0.78026   0.54381
  Math  -0.80603  -0.44283
  Phys  -0.59683  -0.71379
  Chem  -0.72185  -0.63568


Principal components...
    ID   comp.1    comp.2
     1  1.14263  -0.31126
     2  0.85614   0.13922
     3  1.39156  -0.16393
           .
           .
           .
    18 -0.16398   0.01236
    19  0.25673   0.31686
    20  0.11365   0.95938
```

```
Varimax rotation was applied...

 Pattern for the Varimax rotated components =
            comp.1    comp.2
    Jpn  -0.04003   0.96874
    Eng  -0.21177   0.90061
   Hist  -0.11008   0.94468
   Math  -0.86587   0.30993
   Phys  -0.93005  -0.02632
   Chem  -0.95445   0.11907

 Sorted Pattern for Varimax rotation =
            comp.1    comp.2
   Chem  -0.95445   0.11907
   Phys  -0.93005  -0.02632
   Math  -0.86587   0.30993
    Jpn  -0.04003   0.96874
   Hist  -0.11008   0.94468
    Eng  -0.21177   0.90061

Components for Varimax...
    ID    comp.1    comp.2
     1   0.52441  -1.06183
     2   0.67176  -0.54871
     3   0.79973  -1.15054
              .
              .
              .
    18  -0.09944   0.13097
    19   0.40742   0.01783
    20   0.79363   0.55088

Promax rotation was applied...

Pattern for the Promax rotated components =
            comp.1    comp.2
    Jpn   0.08941   0.98905
    Eng  -0.09616   0.89541
   Hist   0.01422   0.95468
   Math  -0.84757   0.19873
```



```
   Phys  -0.95880  -0.15534
   Chem  -0.96427  -0.00944

Correlation matrix of the components...
            comp.1    comp.2
 comp.1   1.00000  -0.26097
 comp.2  -0.26097   1.00000

Sorted Pattern for Promax rotation =
   Chem  -0.96427  -0.00944
   Phys  -0.95880  -0.15534
   Math  -0.84757   0.19873
    Jpn   0.08941   0.98905
   Hist   0.01422   0.95468
    Eng  -0.09616   0.89541

Components for Promax...
    ID    comp.1    comp.2
     1   0.66116  -1.12102
     2   0.73886  -0.63142
     3   0.94585  -1.24479
              .
              .
              .
    18  -0.11600   0.14279
    19   0.40142  -0.03532
    20   0.71319   0.44297
```

参 考 文 献

Okamoto, Y. (2006). A Justification of Rotation in Principal Component Analysis : Projective viewpoint of PCA. *Japan Women's University: Faculty of Integrated Arts and Social Sciences journal*, 17, 59–71. Retrieved April 2018, from https://ci.nii.ac.jp/naid/110006223025.