

# 数量化 I 類

青木繁伸

2020 年 3 月 17 日

## 1 目的

数量化 I 類による分析を行う。

カテゴリー変数をダミー変数に変換すれば、R に元から備わっている `lm` 関数で分析することにより本質的に同じ解を得ることができる。

## 2 使用法

```
import sys
sys.path.append("statlib")
from multi import qt1
qt1(dat, verbose=True, mute=True)
```

結果の図示

```
import sys
sys.path.append("statlib")
from multi import qt1_plot
qt1_plot(obj, type="p", color="blue", alpha=0.3)
```

### 2.1 引数

<code>dat</code>	最後の列が従属変数，それ以外は全て独立変数であるデータフレーム（二次元配列または二重リストでもよいが，結果に変数名を出せない）
<code>verbose</code>	必要最小限のプリント出力をする
<code>mute</code>	デフォルト（ <code>True</code> ）では予測値はプリント表示しない。プリント表示が必要な場合は <code>False</code> にする。
<code>obj</code>	<code>qt1()</code> の戻り値
<code>type</code>	<code>type="p"</code> （デフォルト）ならば，実測値と予測値の図示。 <code>type="c"</code> ならば，カテゴリースコアの図示。
<code>color</code>	描画色（デフォルトは <code>"blue"</code> ）
<code>alpha</code>	アルファチャンネル（デフォルトは <code>0.3</code> ）

## 2.2 戻り値の名前

"x"	ダミー変数行列と従属変数
"coefficients"	カテゴリースコア
"partial"	偏相関係数
"observed"	実測値
"predicted"	予測値
"name"	カテゴリー名

## 3 使用例

```
import pandas as pd

dat = pd.read_csv("data/qt1.csv")

import sys
sys.path.append("statlib")
from multi import qt1

a = qt1(dat)
```

Quantification theory type I

	Category score
X1-0	-4.874791
X1-1	4.874791
X2-0	1.655897
X2-1	-1.509151
X2-2	-0.142430
X3-0	-4.513890
X3-1	4.336874
X4-0	-2.998508
X4-1	2.998508
X5-0	6.104664
X5-1	-0.552423
X5-2	-5.193143
X6-0	-5.645618
X6-1	-0.475560
X6-2	5.941144
X7-0	-1.304480
X7-1	1.304480

X8-0	1.471549
X8-1	0.186583
X8-2	-1.609363
X9-0	-0.869866
X9-1	0.869866
X10-0	-3.233275
X10-1	1.643099
X10-2	1.543407
X11-0	-1.877999
X11-1	-1.896307
X11-2	3.664374
X12-0	-3.434981
X12-1	0.900074
X12-2	2.460351
constant	44.998000

Correlation coefficient

	X1	X2	X3	X4	X5	X6	X7 \
X1	1.000000	0.255448	0.100020	0.120000	0.147229	0.157135	0.080000
X2	0.255448	1.000000	0.063618	0.150592	-0.099774	-0.095687	-0.087009
X3	0.100020	0.063618	1.000000	0.180036	-0.013802	0.192662	-0.180036
X4	0.120000	0.150592	0.180036	1.000000	0.184905	-0.178888	0.040000
X5	0.147229	-0.099774	-0.013802	0.184905	1.000000	0.212676	0.225222
X6	0.157135	-0.095687	0.192662	-0.178888	0.212676	1.000000	0.276393
X7	0.080000	-0.087009	-0.180036	0.040000	0.225222	0.276393	1.000000
X8	0.115423	-0.069355	-0.266430	-0.026262	0.028569	0.124111	-0.002075
X9	0.080000	-0.109318	-0.060012	-0.120000	0.285699	0.232886	0.120000
X10	-0.146891	0.026580	-0.181570	0.146012	0.158471	-0.119059	0.022365
X11	0.210521	0.300963	0.280728	0.041798	-0.243139	-0.078928	-0.168515
X12	0.052149	-0.020803	0.144215	0.253780	0.049017	-0.061830	-0.089689
Y	0.527076	0.185881	0.419431	0.361966	0.463091	0.441046	0.178381

	X8	X9	X10	X11	X12	Y
X1	0.115423	0.080000	-0.146891	0.210521	0.052149	0.527076
X2	-0.069355	-0.109318	0.026580	0.300963	-0.020803	0.185881
X3	-0.266430	-0.060012	-0.181570	0.280728	0.144215	0.419431
X4	-0.026262	-0.120000	0.146012	0.041798	0.253780	0.361966
X5	0.028569	0.285699	0.158471	-0.243139	0.049017	0.463091
X6	0.124111	0.232886	-0.119059	-0.078928	-0.061830	0.441046
X7	-0.002075	0.120000	0.022365	-0.168515	-0.089689	0.178381
X8	1.000000	-0.010137	0.110739	-0.009576	0.098336	0.108286
X9	-0.010137	1.000000	0.109189	-0.168375	-0.029896	0.181504

X10	0.110739	0.109189	1.000000	0.188254	0.117103	0.158577
X11	-0.009576	-0.168375	0.188254	1.000000	0.050169	0.266153
X12	0.098336	-0.029896	0.117103	0.050169	1.000000	0.288285
Y	0.108286	0.181504	0.158577	0.266153	0.288285	1.000000

Partial correlation coefficient

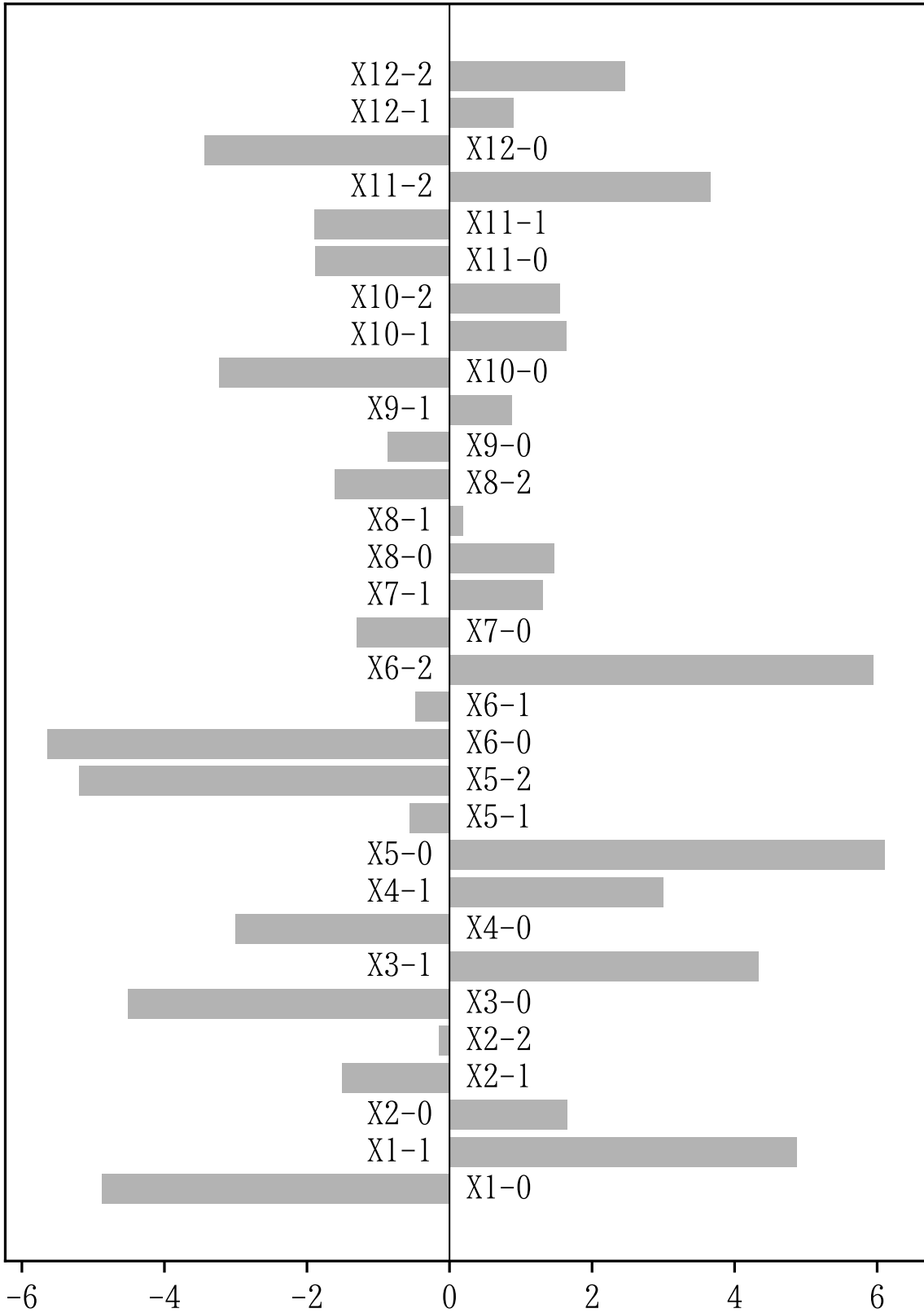
	partial corr. coeff.	t value	p value
X1	0.524609	5.747657	1.312711e-07
X2	0.167638	1.586068	1.163521e-01
X3	0.455754	4.775827	7.188569e-06
X4	0.353229	3.521723	6.852845e-04
X5	0.496526	5.335451	7.470013e-07
X6	0.497940	5.355642	6.869082e-07
X7	0.167603	1.585727	1.164294e-01
X8	0.164755	1.558019	1.228591e-01
X9	0.114161	1.071830	2.867608e-01
X10	0.272876	2.645621	9.675575e-03
X11	0.295421	2.884232	4.944678e-03
X12	0.322262	3.175260	2.071085e-03

### 3.1 カテゴリースコアの図示

```
from multi import qt1_plot

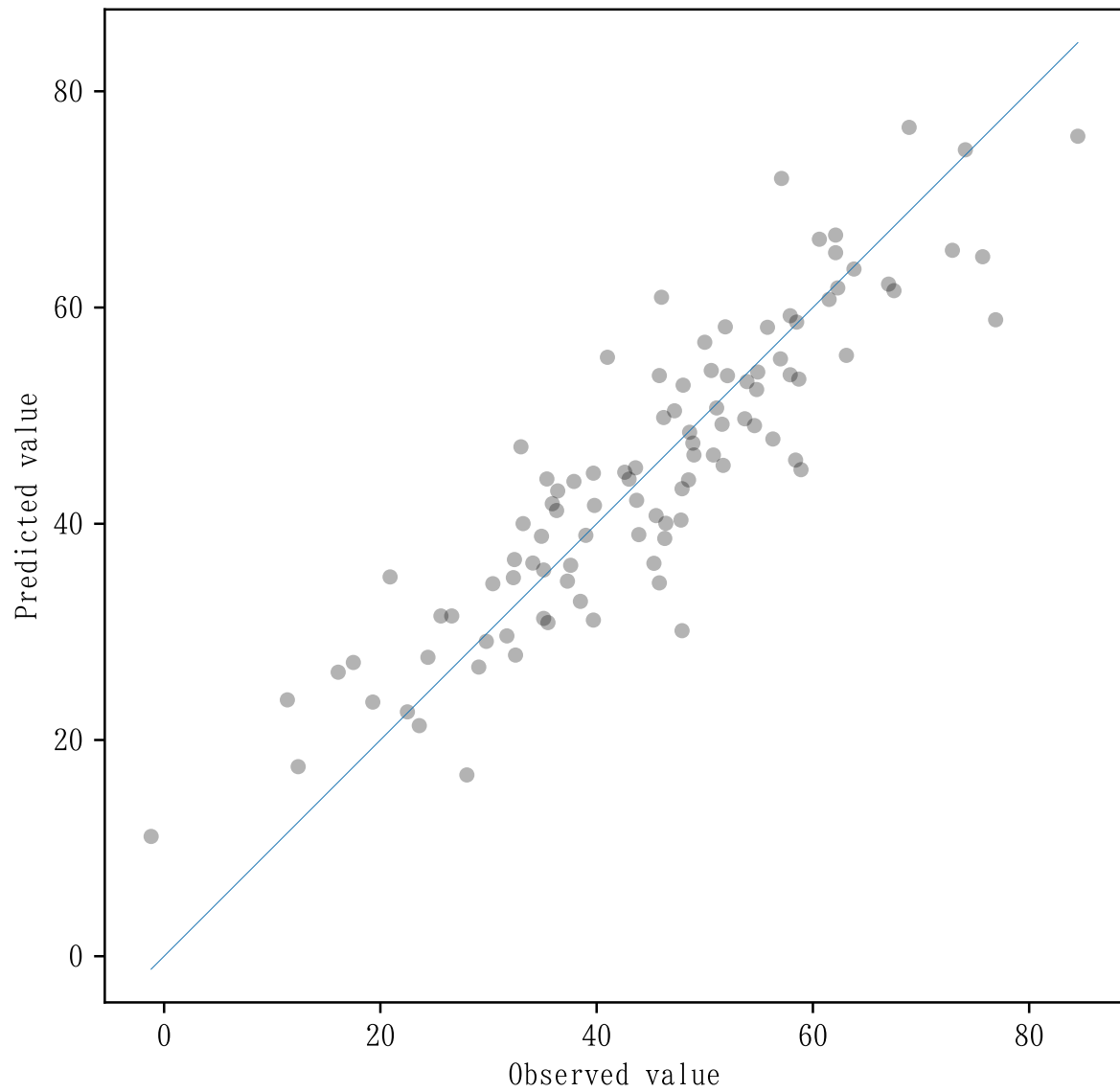
qt1_plot(a, type="c")
```

# Category score



### 3.2 予測値の図示

```
qt1_plot(a, type="p")
```



```
import pandas as pd

x1 =[1,3,1,1,2,1,2,1,3,1]
x2 =[2,2,2,1,3,3,2,1,2,1]
x3 =[2,2,2,1,2,3,2,1,2,2]
y =[6837,7397,7195,6710,6670,6279,6601,4929,5471,6164]
dat = pd.DataFrame({"x1":x1,"x2":x2,"x3":x3,"y":y})

import sys
```

```

sys.path.append("statlib")
from multi import qt1

a = qt1(dat)

```

Quantification theory type I

	Category score
x1-1	199.4
x1-2	-215.6
x1-3	-382.6
x2-1	-610.2
x2-2	241.8
x2-3	310.8
x3-1	-195.0
x3-2	149.5
x3-3	-656.5
constant	6425.3

Correlation coefficient

	x1	x2	x3	y
x1	1.000000	-0.510664	-0.463184	-0.103277
x2	-0.510664	1.000000	0.164788	0.440587
x3	-0.463184	0.164788	1.000000	0.290518
y	-0.103277	0.440587	0.290518	1.000000

Partial correlation coefficient

	partial corr. coeff.	t value	p value
x1	0.308748	0.795121	0.456835
x2	0.500942	1.417768	0.206039
x3	0.358395	0.940354	0.383335