

数量化 I 類

青木繁伸

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

```
from qt1 import qt1
qt1(dat, verbose=True)
結果の図示
from qt1 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>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>	アルファチャネル（デフォルトは <code>0.3</code> ）

2.2 戻り値の名前

<code>"x"</code> <code>"coefficients"</code>	ダミー変数行列と従属変数
	カテゴリースコア
<code>"partial"</code>	偏相関係数
<code>"observed"</code>	実測値
<code>"predicted"</code>	予測値
<code>"name"</code>	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 qt1 import qt1

a = qt1(dat)
```

数量化 I 類

カテゴリースコア

カテゴリースコア

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
定数項	44.998000

相関係数

	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

偏相関係数

	偏相関係数	t 値	p 値
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

予測値			
	観察値	予測値	残差
0	75.7	64.697898	11.002102
1	32.5	27.853704	4.646296
2	12.4	17.532317	-5.132317
3	72.9	65.292440	7.607560
4	11.4	23.711104	-12.311104
5	52.1	53.698466	-1.598466
6	67.5	61.559505	5.940495
7	45.8	34.531580	11.268420
8	32.4	36.688841	-4.288841
9	54.8	52.415375	2.384625
10	37.3	34.692662	2.607338
11	33.2	40.012090	-6.812090
12	57.9	53.778235	4.121765
13	34.1	36.366704	-2.266704
14	47.9	30.110193	17.789807
15	32.3	35.014313	-2.714313
16	-1.2	11.079673	-12.279673
17	57.0	55.249488	1.750512
18	37.9	43.912977	-6.012977
19	48.5	44.056680	4.443320
20	43.7	42.168171	1.531829
21	57.9	59.240740	-1.340740
22	51.1	50.711055	0.388945
23	36.3	41.223864	-4.923864
24	50.0	56.786444	-6.786444
25	51.9	58.212754	-6.312754
26	35.9	41.856806	-5.956806
27	36.4	43.037749	-6.637749
28	53.7	49.708729	3.991271
29	24.4	27.647730	-3.247730
..
70	19.3	23.506976	-4.206976
71	45.3	36.342587	8.957413
72	76.9	58.867433	18.032567
73	39.7	44.679294	-4.979294
74	29.8	29.123403	0.676597
75	62.3	61.813892	0.486108
76	48.6	48.463008	0.136992
77	74.1	74.588841	-0.488841
78	45.8	53.705677	-7.905677
79	46.0	60.950551	-14.950551
80	56.3	47.838681	8.461319

```
81 42.6 44.766729 -2.166729
82 35.5 30.858341 4.641659
83 28.0 16.766681 11.233319
84 43.0 44.123079 -1.123079
85 41.0 55.397188 -14.397188
86 47.9 43.238390 4.661610
87 22.5 22.600609 -0.100609
88 54.6 49.077223 5.522777
89 47.2 50.447641 -3.247641
90 35.4 44.142117 -8.742117
91 84.5 75.842983 8.657017
92 68.9 76.663223 -7.763223
93 53.9 53.149361 0.750639
94 58.7 53.380031 5.319969
95 50.6 54.175422 -3.575422
96 25.6 31.470851 -5.870851
97 46.4 40.042570 6.357430
98 34.9 38.842069 -3.942069
99 60.6 66.315886 -5.715886
```

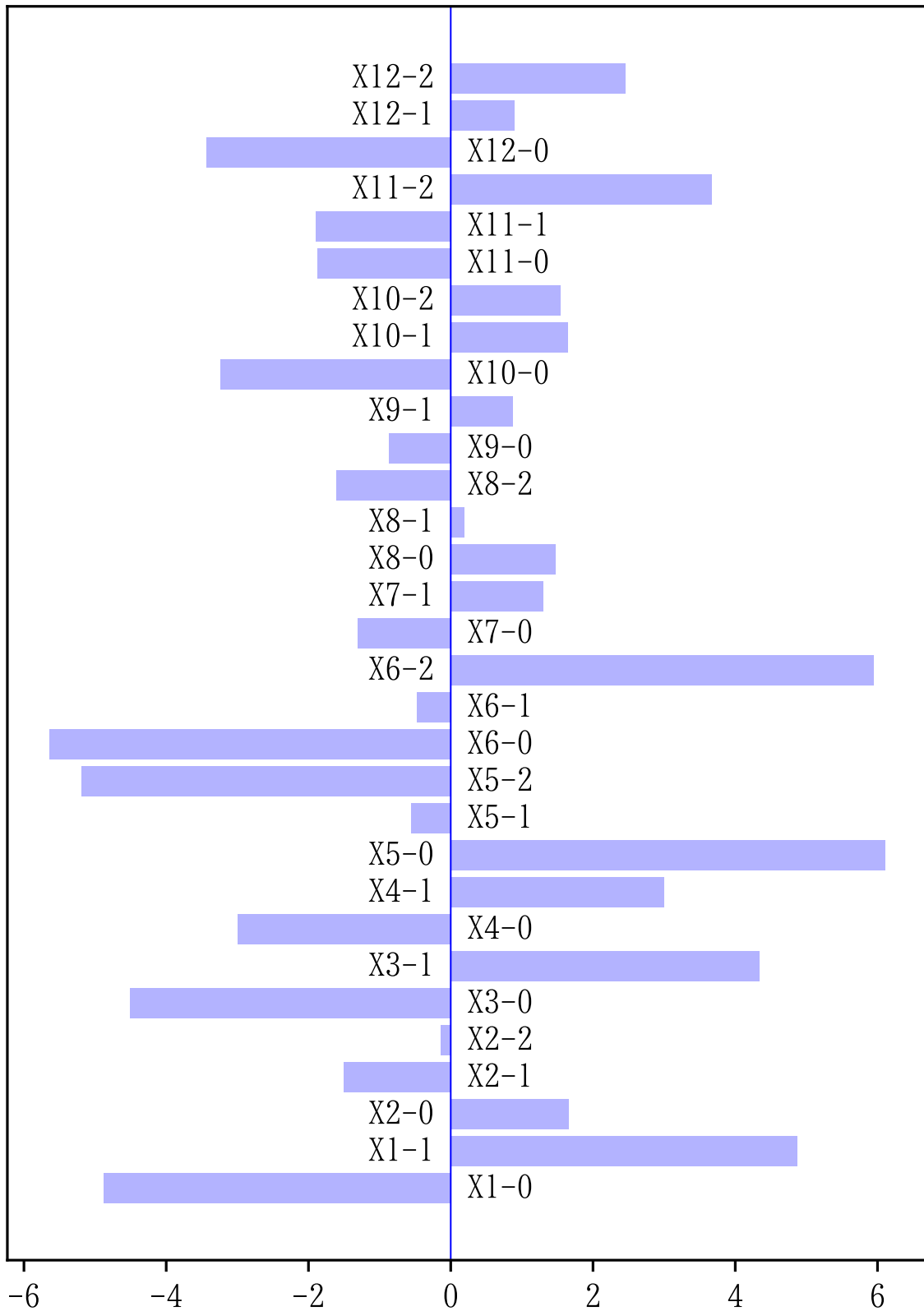
[100 rows x 3 columns]

3.1 カテゴリースコアの図示

```
from qt1 import qt1_plot

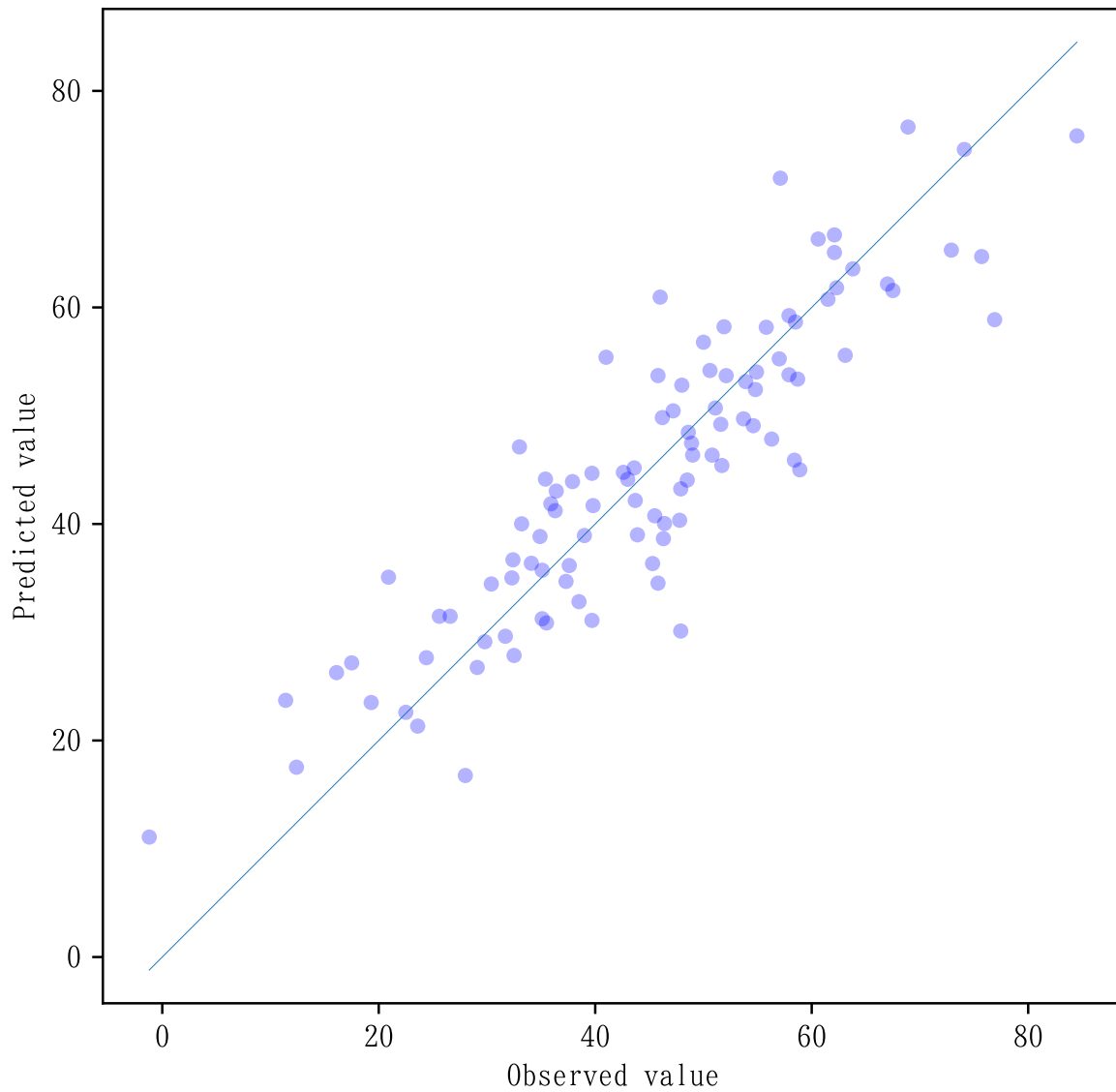
qt1_plot(a, type="c")
```

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 qt1 import qt1

a = qt1(dat)

```

数量化 I 類

カテゴリースコア

カテゴリースコア

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
定数項	6425.3

相関係数

	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

偏相関係数

	偏相関係数	t 値	p 値
x1	0.308748	0.795121	0.456835
x2	0.500942	1.417768	0.206039
x3	0.358395	0.940354	0.383335

予測値

	観察値	予測値	残差
0	6837.0	7016.0	-1.790000e+02
1	7397.0	6434.0	9.630000e+02
2	7195.0	7016.0	1.790000e+02
3	6710.0	5819.5	8.905000e+02
4	6670.0	6670.0	9.094947e-13
5	6279.0	6279.0	0.000000e+00
6	6601.0	6601.0	-9.094947e-13
7	4929.0	5819.5	-8.905000e+02
8	5471.0	6434.0	-9.630000e+02
9	6164.0	6164.0	-9.094947e-13