

クロス集計表・分布表の双対尺度法

青木繁伸

1 目的

二次元クロス集計表から、双対尺度法による解析を行う。

RのMASSライブラリに入っている `corresp` 関数と同じであるが、若干親切な出力をする。

2 使用法

```
import sys
sys.path.append("statlib")
from dual import dual
dual(tbl, max\_axis=5, verbose=True)
```

行スコアと列スコアのバイプロットを描く

```
import sys
sys.path.append("statlib")
from dual import dual_plot
dual_plot(obj, weighted=False, ax1=1, ax2=2, color="blue", color2="red", alpha=0.5):
```

2.1 引数

<code>tbl</code>	二次元クロス表
<code>max_axis</code>	解の個数を制限する (デフォルトは 5)
<code>verbose</code>	必要最小限のプリント出力をする
<code>obj</code>	<code>dual()</code> の戻り値
<code>weighted</code>	<code>True</code> を指定すれば重み付きの行・列スコアをプロットする
<code>ax1</code>	横軸にとる解の番号
<code>ax2</code>	縦軸にとる解の番号
<code>color</code>	列スコアに対する点とテキストの色
<code>color2</code>	行スコアに対する点とテキストの色
<code>alpha</code>	アルファチャンネル (デフォルトは 0.5)

2.2 戻り値の名前

<code>"result"</code>	結果表
<code>"rs"</code>	行スコア
<code>"cs"</code>	列スコア
<code>"wrs"</code>	重み付き行スコア

3 使用例

```
tbl = [[2, 3, 5, 6],
        [5, 1, 7, 5],
        [5, 3, 4, 3]]

import sys
sys.path.append("statlib")
from dual import dual

a = dual(tbl)
```

Summary

	Axis 1	Axis 2
eta square	0.049054	0.037657
correlation	0.221481	0.194053
contribution	56.572123	43.427877
cumulative contribution	56.572123	100.000000
Chi square value	2.263408	1.727273
degrees of freedom	4.000000	2.000000
p value	0.687439	0.421626

Row score

	Axis 1	Axis 2
R1	-1.435585	0.039955
R2	0.665093	-1.131315
R3	0.733178	1.314959

Column score

	Axis 1	Axis 2
C1	1.550240	0.428633
C2	-0.930182	2.159520
C3	0.115823	-0.792178
C4	-0.996055	-0.541813

Weighted row score

	Axis 1	Axis 2
R1	-1.435585	0.039955
R2	0.665093	-1.131315
R3	0.733178	1.314959

Weighted column score

	Axis 1	Axis 2
--	--------	--------

```

C1 1.550240 0.428633
C2 -0.930182 2.159520
C3 0.115823 -0.792178
C4 -0.996055 -0.541813

```

```

import pandas as pd

dat = pd.read_csv("data/ooyama-okamoto.csv", index_col=0)

import sys
sys.path.append("statlib")
from dual import dual

a = dual(dat)

```

Summary

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
eta square	0.795495	0.696514	0.580795	0.448228	0.314787
correlation	0.891905	0.834574	0.762099	0.669498	0.561059
contribution	22.773618	19.939963	16.627136	12.831968	9.011788
cumulative contribution	22.773618	42.713581	59.340718	72.172686	81.184474
Chi square value	45.234160	33.983963	24.777771	16.946669	10.773718
degrees of freedom	25.000000	23.000000	21.000000	19.000000	17.000000
p value	0.007863	0.065390	0.256917	0.593481	0.868082

Row score

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
anger	-0.647286	0.619050	-0.159263	-0.759045	-0.271959
jealousy	-0.584775	0.936736	-0.258833	-1.549712	1.393995
sin	-1.797653	-0.826087	0.310674	1.691387	-1.630901
eternity	0.980942	-1.273413	-1.229405	-0.626209	-0.910561
happy	0.544964	1.522066	0.722808	0.317164	-0.848015
lonely	-0.784685	-0.953140	0.355571	0.319553	0.153121
calm	1.133635	-1.350364	0.322556	-0.099561	0.635229
nostalgia	1.149231	-0.789359	2.106888	-0.206966	0.337644
home	0.544964	1.522066	0.722808	0.317164	-0.848015
love	0.084096	1.368666	-0.253590	-0.373234	-0.437080
purity	0.573125	-0.544725	-1.749380	-1.096666	-1.036131
dream	1.089055	0.527654	-1.510201	2.513161	1.549823
anxiety	-1.476188	-0.271580	0.081475	-0.227142	1.772751
fear	-1.192502	-0.224452	-0.164404	-0.150905	0.027551

Column score

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
red	-0.396307	0.516497	-0.678513	-1.173883	-0.115362
orange	-0.013014	1.430331	0.203105	-0.611701	-0.360417

black	-1.322633	-0.396899	0.111285	0.261046	0.018024
purple	-1.155371	0.398500	-0.116362	-1.327004	2.822117
gray	-1.471857	-0.681563	0.191352	0.609745	0.143711
violet	-2.015520	-0.989830	0.407655	2.526351	-2.906829
white	0.871206	-1.089261	-1.954329	-1.286692	-1.734838
grnble	1.058620	-0.791076	-1.366761	0.257926	0.106210
pink	0.634339	1.479932	-0.104375	1.035946	-0.259905
yorange	0.836845	0.900569	1.553824	0.212777	-0.807037
blue	0.694895	-1.307935	0.510304	-0.228971	0.095994
green	1.279769	-1.281925	1.593916	-0.228923	0.866998
ygrn	1.288513	-0.945822	2.764585	-0.309136	0.601799
yellow	1.221044	0.632243	-1.981633	3.753798	2.762320

Weighted row score

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
anger	-0.647286	0.619050	-0.159263	-0.759045	-0.271959
jealousy	-0.584775	0.936736	-0.258833	-1.549712	1.393995
sin	-1.797653	-0.826087	0.310674	1.691387	-1.630901
eternity	0.980942	-1.273413	-1.229405	-0.626209	-0.910561
happy	0.544964	1.522066	0.722808	0.317164	-0.848015
lonely	-0.784685	-0.953140	0.355571	0.319553	0.153121
calm	1.133635	-1.350364	0.322556	-0.099561	0.635229
nostalgia	1.149231	-0.789359	2.106888	-0.206966	0.337644
home	0.544964	1.522066	0.722808	0.317164	-0.848015
love	0.084096	1.368666	-0.253590	-0.373234	-0.437080
purity	0.573125	-0.544725	-1.749380	-1.096666	-1.036131
dream	1.089055	0.527654	-1.510201	2.513161	1.549823
anxiety	-1.476188	-0.271580	0.081475	-0.227142	1.772751
fear	-1.192502	-0.224452	-0.164404	-0.150905	0.027551

Weighted column score

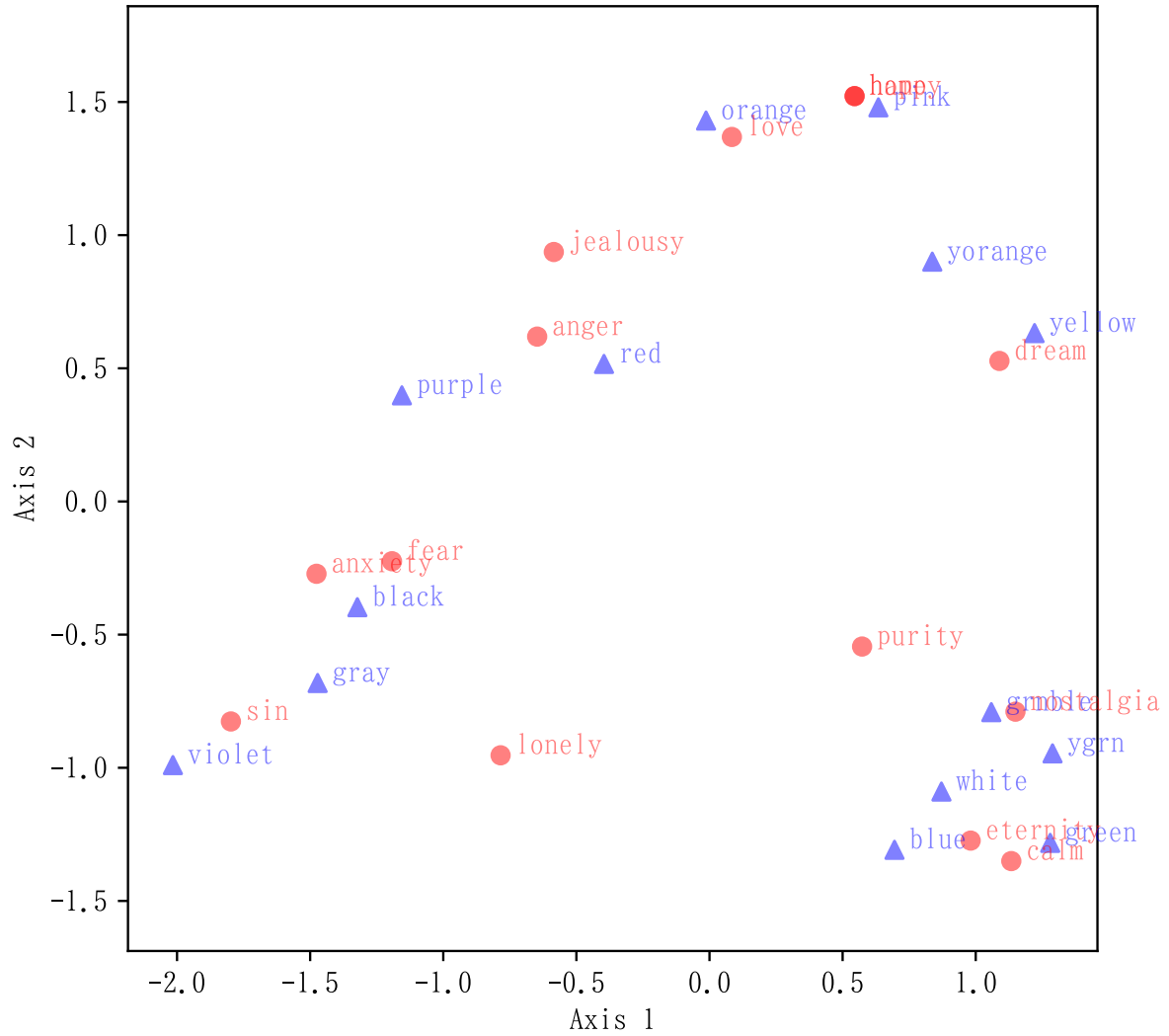
	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
red	-0.396307	0.516497	-0.678513	-1.173883	-0.115362
orange	-0.013014	1.430331	0.203105	-0.611701	-0.360417
black	-1.322633	-0.396899	0.111285	0.261046	0.018024
purple	-1.155371	0.398500	-0.116362	-1.327004	2.822117
gray	-1.471857	-0.681563	0.191352	0.609745	0.143711
violet	-2.015520	-0.989830	0.407655	2.526351	-2.906829
white	0.871206	-1.089261	-1.954329	-1.286692	-1.734838
grnble	1.058620	-0.791076	-1.366761	0.257926	0.106210
pink	0.634339	1.479932	-0.104375	1.035946	-0.259905
yorange	0.836845	0.900569	1.553824	0.212777	-0.807037
blue	0.694895	-1.307935	0.510304	-0.228971	0.095994
green	1.279769	-1.281925	1.593916	-0.228923	0.866998
ygrn	1.288513	-0.945822	2.764585	-0.309136	0.601799

yellow 1.221044 0.632243 -1.981633 3.753798 2.762320

```
import sys
sys.path.append("statlib")
from dual import dual_plot

dual_plot(a)
```

Dual Scale Analysis (non-weighted scores)



```
dual_plot(a, weighted=True)
```

Dual Scale Analysis (weighted scores)

