

判別分析

統計数理研究所(名誉教授) 鈴木義一郎

§ 1 判別分析モデルの平均ベクトルと分散共分散行列の推定

```
damean=: (mean=:+/%#)"1&> NB. ボックス形で与えられたデータの平均ベクトル
dav=: (: (+/ .* |:)) (-mean)"1 NB. 行と列の変動に関する偏差平方和を与える
davarm=: (: (+/dav&>)%[:+#@|:)&> NB. 分散共分散行列を与える関数
```

```
MIN2=. 2.1 1.5 2.9 3 3.3 3.7 4.7 4.2 4.8 4.9 3 2.9 1.1 1.8
MIN2=:MIN2, 2 3.7 5 3.3 0.5 0.9 0.5 _1.4 _0.4 2.3 3.9 8.4 7.2 6.7
MIN3=:6.7 4.6 5.2 4 5 6.7 13.4 9.1 6.6 4.6 7.7 6.1 4.6 9.2 6.2 4.3
MIN3=:MIN3, 3.5 5.6 6.9 8 6.3 8.4 5.4 4.2 3 10.4 8.4 6.3 12 10.4 7.9
MAX2=:9.9 8.6 9.1 8.4 10.1 9.4 11.7 10.2 12.5 12.3 12.6 10.4 9.5 10.8
MAX2=:MAX2, 11.2 9.3 11.8 10.9 9 12.1 13.9 4.6 9.6 13.1 16.1 18 13.4 13.9
MAX3=:10+10 1.3 _2.3 1.4 5.2 5.6 11.9 5.7 4.6 4.3 10.2 4.7 3.7 8.4 1.2 _3.6
MAX3=:MAX3, 10+2.4 3.4 4.4 3 5.7 3.1 _0.4 0.1 5.9 7.5 1.9 7.2 7.4 15.1 4.6
A=:MIN2, :MAX2 [ B=:MIN3, :MAX3
NB. 東京地区の平成9年の2月と3月の最低気温と最高気温のデータ
```

<pre>]M=:damean A;B 3.08929 11.1571 6.79677 14.6323</pre>	<pre>]MO=: -: @+ /damean A;B 4.94303 12.8947</pre>	<p>1行目、2行目がそれぞれA、Bの平均ベクトルMで、それらの中点を与えている。</p>				
<pre>dav L:0 A;B</pre> <table border="1" data-bbox="231 1731 721 1809"> <tr> <td>134.867 101.817</td> <td>187.47 190.273</td> </tr> <tr> <td>101.817 184.009</td> <td>190.273 479.068</td> </tr> </table>		134.867 101.817	187.47 190.273	101.817 184.009	190.273 479.068	<p>AとBに「dav」を実行した結果をボックス形で出力している。 、これらの合計をQに挿入して表示している。</p>
134.867 101.817	187.47 190.273					
101.817 184.009	190.273 479.068					
<pre>]N=:#@ :)&> A;B 28 31</pre>	<pre>]Q=:+ /dav&> A;B 322.336 292.09</pre>	<p>AとBそれぞれのデータ数と分散共分散行列の合計をQに挿入している。</p>				

	292.09 663.076	
Q % +/N 5.46333 4.95068 4.95068 11.2386]V=:davarm A;B 5.46333 4.95068 4.95068 11.2386	A、B 2種類のデータの分散共分散行列の加重平均を出力。「davarm」は2(多)次元データの対の分散共分散行列を出力する。

§ 2 線形判別関数

```
dacoef=:[:-/damean %."1 2 davarm NB. 線形判別関数の係数を与える関数
dapoint=:dacoef*[::@+/damean NB. 線形判別関数の判別点を与える関数
davalue=:3 :0
NB. 線形判別関数の値を与える関数
(p=. dapoint y);c=. dacoef y
(+/(c*>{. y)-p);+/(c*>{:y)-p
)
```

-/ M %. V _0.562971 0.765649]C=:dacoef A;B _0.663111 _0.0171071	線形判別関数の係数を求めてCに挿入し表示している。
]MO=::@+/ damean A;B 4.94303 12.8947 C * MO _3.27778 _0.220591]P=:dapoint A;B 3.27778 _0.220591		AとBの平均ベクトルの中点を求めて、MOに挿入し表示している。 MOに線形判別関数の係数Cを掛けている。 上と同じ結果で、2つのグループに判別するときの判別点を与える。
]5{."1 DA=:C * A _1.39253 _0.994667 _1.92302 _1.98933 _2.18827 _0.169361 _0.147121 _0.155675 _0.1437 _0.172782]5{."1 DB=:C * B _4.44285 _3.05031 _3.44818 _2.65245 _3.31556 0.342143 0.193311 0.131725 0.195021 0.260028		
5{."1 DA-P 1.88525 2.28311 1.35476 1.28845 1.08951		

PH=:PH;0 5 14 2 15 9 11 62 324 5 NB. 6/25	ウクライナ:スイス
IA=:1 2 28 2 17 6 5 41 237 4 IA=:IA;0 2 19 2 26 4 4 59 387 3 NB. 6/26	ブラジル:ガーナ
US=:0 6 24 1 20 2 8 45 269 2 US=:US;0 5 21 0 24 6 6 55 108 4 NB. 6/26	フランス:スペイン
BG=:3 3 28 5 18 10 1 48 321 1 BG=:BG;0 4 23 4 24 7 11 52 348 3 NB. 6/27	
FS=:3 5 31 8 23 5 4 39 190 1 FS=:FS;1 8 31 2 29 2 5 61 371 2 NB. 6/27	

【準々決勝】

DA=:1 4 45 3 23 5 5 42 246 2 DA=:DA;1 6 26 3 32 5 7 58 384 2 NB. 6/31 IU=:3 1 33 2 15 7 3 41 196 2 IU=:IU;0 3 17 2 31 7 6 59 285 3 NB. 6/31 PE=:0.5 4 21 3 10 9 11 57 507 2 PE=:PE;0 6 13 0 21 4 5 43 335 3 NB. 7/1 FB=:1 7 24 5 17 5 4 45 258 3 FB=:FB;0 5 22 2 22 1 6 55 312 3 NB. 7/1	ドイツ:アルゼンチン イタリア:ウクライナ ポルトガル:イングランド フランス:ブラジル
---	---

【準決勝】

ID=:2 12 23 11 19 10 5 57 396 2 ID=:ID;0 4 30 2 21 2 11 43 314 5 NB. 7/4 FP=:1 3 22 0 11 4 1 41 269 1 FP=:FP;0 8 11 4 18 5 7 59 283 3 NB. 7/5	イタリア:ドイツ フランス:ポルトガル
---	----------------------------

【3位決定戦】

DP=:3 2 18 1 15 5 7 43 208 2 DP=:DP;0 7 16 3 15 8 5 57 363 5 NB. 7/8	【決勝】 IF=:1 5 26 4 17 3 2 55 343 1 IF=:IF;1 7 21 2 24 5 8 45 393 3 NB. 7/9
---	---

【決勝トーナメントデータの判別分析】

\$A1=: :}."1>{."1 TT1 9 8 \$B1=: :}:"1>{."1 TT1 9 8	\$A2=: :}."1>{."1 TT2 9 8 \$B2=: :}:"1>{."1 TT2 9 8	\$A=: :}."1>{."1 TT 9 16 \$B=: :}:"1>{."1 TT 9 16												
TD1=:davalue A1;B1 .. (>0:)L:0 TD1 <table border="1" style="width: 100%; text-align: center;"> <tr><td>1 1 1 1 1 1 1</td></tr> <tr><td>1</td></tr> <tr><td>0 0 0 0 0 0 0</td></tr> <tr><td>0</td></tr> </table>	1 1 1 1 1 1 1	1	0 0 0 0 0 0 0	0	TD2=:davalue A2;B2 .. (>0:)L:0 TD2 <table border="1" style="width: 100%; text-align: center;"> <tr><td>1 1 1 1 1 1 1</td></tr> <tr><td>1</td></tr> <tr><td>0 0 0 0 0 0 0</td></tr> <tr><td>0</td></tr> </table>	1 1 1 1 1 1 1	1	0 0 0 0 0 0 0	0	TD2=:davalue A;B .. (>0:)L:0 TD <table border="1" style="width: 100%; text-align: center;"> <tr><td>1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1</td></tr> <tr><td>1</td></tr> <tr><td>0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0</td></tr> <tr><td>0</td></tr> </table>	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0
1 1 1 1 1 1 1														
1														
0 0 0 0 0 0 0														
0														
1 1 1 1 1 1 1														
1														
0 0 0 0 0 0 0														
0														
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1														
1														
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0														
0														
6.2":dacoef A1;B1		トーナメント一回戦の判別												

2.43 1.52 _1.14 2.68 _1.13 _2.52 4.94 _0.36 0.11	関数のパラメータ。
6.2":dacoef A2;B2 _7.10 _0.06 _0.92 4.66 _2.07 _1.88 5.49 _0.05 0.28	準々決勝・準決勝・3位決定戦・決勝
6.2":dacoef A;B _1.89 0.65 _0.60 1.93 _1.47 _1.72 3.65 _0.11 0.14	全トーナメントでの判別関数のパラメータ。

ワールドカップデータに関しては JAPLA 9/2006 の鈴木を参照