

主成分分析用データファイルの作成

第 1 ステップで出力された数量化の計算結果の出力ファイルには、最後の数量化以外のものも出力されている。

・
・
・

Lambda...

6.49239 4.56643

Omega...

0.747 -0.331
-0.660 -0.481
-0.087 0.812

Results...

GrpID-0	ID-0	Score-0	Attr.-0	CA-0	CA-1	CB-0	CB-1
GA	A1	45	-1.41077	1.12041	-0.41623	1.12041	-0.41623
GA	A2	51	-0.66130	1.12041	-0.41623	1.12041	-0.41623
GA	A3	39	-1.03604	1.12041	-0.41623	-0.22819	1.78828
GA	A4	67	0.08817	-0.22819	1.78828	1.12041	-0.41623
GA	A5	50	-1.03604	1.12041	-0.41623	1.12041	-0.41623
GB	B1	85	1.21238	-1.15502	-0.70658	1.12041	-0.41623
GB	B2	91	1.58712	-1.15502	-0.70658	1.12041	-0.41623
GB	B3	72	0.46291	-0.22819	1.78828	1.12041	-0.41623
GB	B4	75	0.83765	-1.15502	-0.70658	-0.22819	1.78828
GC	C1	52	0.83765	-1.15502	-0.70658	-1.15502	-0.70658
GC	C2	43	1.58712	-1.15502	-0.70658	-1.15502	-0.70658
GC	C3	65	0.83765	-1.15502	-0.70658	-0.22819	1.78828
GC	C4	38	0.08817	-0.22819	1.78828	-1.15502	-0.70658
GD	D1	15	-1.41077	1.12041	-0.41623	-1.15502	-0.70658
GD	D2	17	-0.66130	1.12041	-0.41623	-1.15502	-0.70658
GD	D3	37	-0.28656	-0.22819	1.78828	-1.15502	-0.70658
GD	D4	35	-1.03604	1.12041	-0.41623	-0.22819	1.78828

主成分分析を行うためには、最後の数量化の部分だけが必要である。これは、出力データファイルを適当なエディタで読み込み、簡単に編集することができる。図 1 は、メモ帳で読み込み、編集した結果を示したものである。元の数量化の出力ファイルを残すため、ファイル名を変えてある。

GrpID-0	ID-0	Score-0	Attr.-0	CA-0	CA-1	CB-0	CB-1
GA	A1	45	-1.41077	1.12041	-0.41623	1.12041	-0.41623
GA	A2	51	-0.66130	1.12041	-0.41623	1.12041	-0.41623
GA	A3	39	-1.03604	1.12041	-0.41623	-0.22819	1.78828
GA	A4	67	0.08817	-0.22819	1.78828	1.12041	-0.41623
GA	A5	50	-1.03604	1.12041	-0.41623	1.12041	-0.41623
GB	B1	85	1.21238	-1.15502	-0.70658	1.12041	-0.41623
GB	B2	91	1.58712	-1.15502	-0.70658	1.12041	-0.41623
GB	B3	72	0.46291	-0.22819	1.78828	1.12041	-0.41623
GB	B4	75	0.83765	-1.15502	-0.70658	-0.22819	1.78828
GC	C1	52	0.83765	-1.15502	-0.70658	-1.15502	-0.70658
GC	C2	43	1.58712	-1.15502	-0.70658	-1.15502	-0.70658
GC	C3	65	0.83765	-1.15502	-0.70658	-0.22819	1.78828
GC	C4	38	0.08817	-0.22819	1.78828	-1.15502	-0.70658
GD	D1	15	-1.41077	1.12041	-0.41623	-1.15502	-0.70658
GD	D2	17	-0.66130	1.12041	-0.41623	-1.15502	-0.70658
GD	D3	37	-0.28656	-0.22819	1.78828	-1.15502	-0.70658
GD	D4	35	-1.03604	1.12041	-0.41623	-0.22819	1.78828

図 1

図 1 のファイルから、変数 GrpID-0 と Score-0 のデータを除くとする。メモ帳で列単位でデータを削除することは煩わしいので、図 1 のデータを Excel で読み込み、Excel で列単位の処理を行う。

まず、Excel でファイルを開くときに、ファイルの種類を「テキストファイル (*.prn,*.txt,*.csv)」を選ぶ

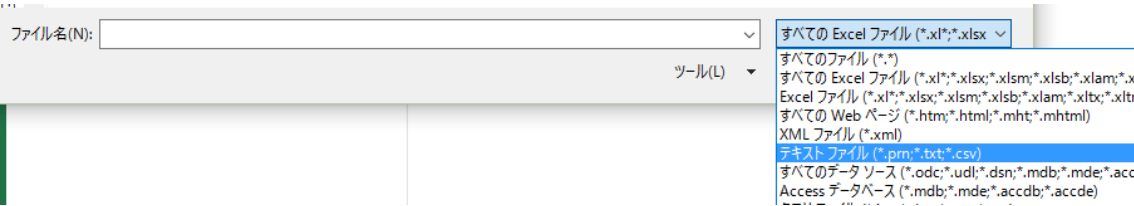


図 2

図 1 のファイル名（拡張子が、.txt のもの）が表示される（図 3）

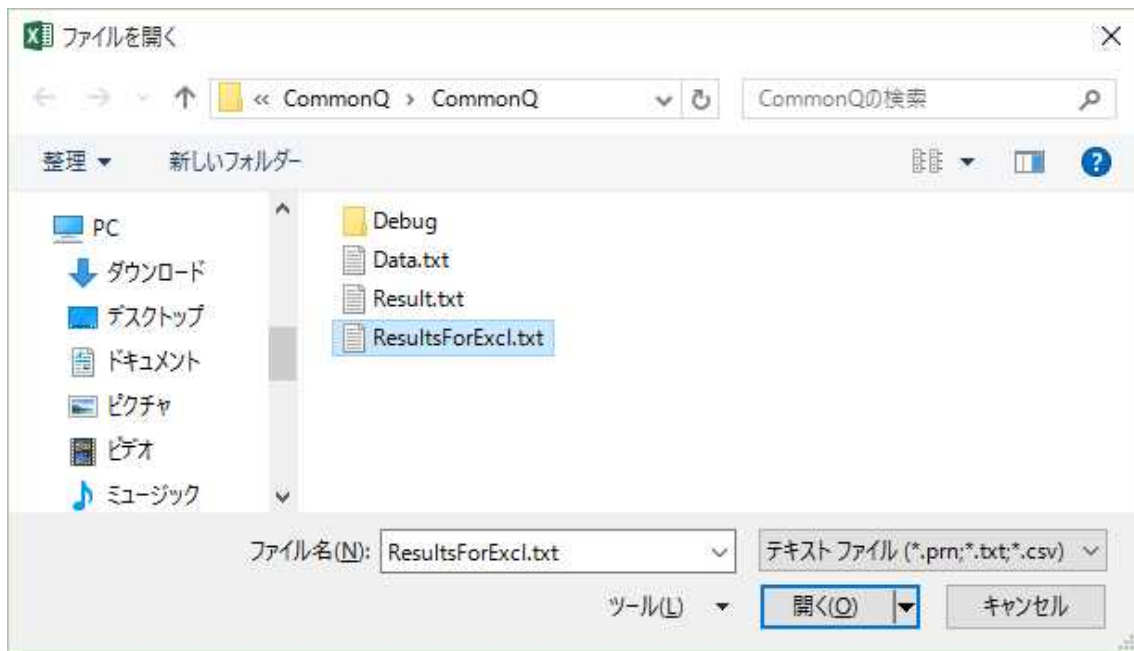


図 3

「ファイル名」を設定して、「開く」ボタンをクリックする。
「テキスト ファイル ウィザード-1/3」が表示される（図 4）。

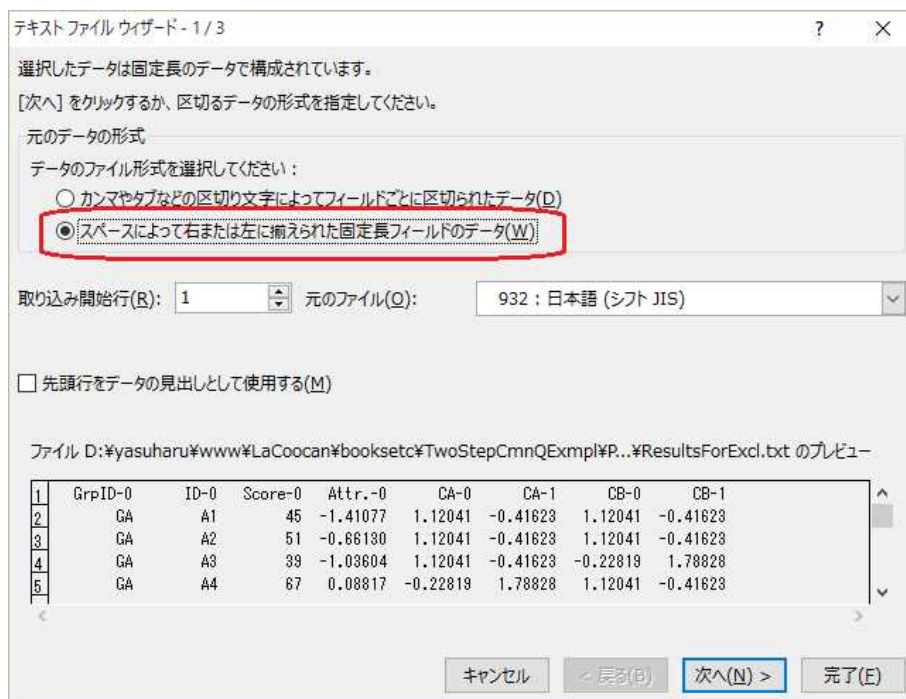


図 4

図 1 のファイルは、列ごとに整列されているので「スペースによって右または左に揃えられ

「固定長フィールドのデータ」にチェックを入れる（図4）。「次へ」をクリックすると、次のウィザードのページが表示される（図5）。

テキストファイルウィザード - 2 / 3

フィールドの幅を指定してください。
 [データのプレビュー] ボックスの矢印の位置でフィールドに区切られます。

区切り位置を作成する場合は、その位置でマウスのボタンをクリックします。
 区切り位置を削除する場合は、その矢印をダブルクリックします。
 区切り位置を移動する場合は、矢印を目的の位置までドラッグします。

データのプレビュー(P)

GrpID-0	ID-0	Score-0	Attr.-0	CA-0	CA-1	CB-0	CB-1
GA	A1	45	-1.41077	1.12041	-0.41623	1.12041	-0.41623
GA	A2	51	-0.66130	1.12041	-0.41623	1.12041	-0.41623
GA	A3	39	-1.03604	1.12041	-0.41623	-0.22819	1.78828
GA	A4	67	0.08817	-0.22819	1.78828	1.12041	-0.41623

キャンセル < 戻る(B) 次へ(N) > 完了(E)

図 5

縦線の位置に問題がないことを確認して、「次へ」をクリックする。

テキストファイルウィザード - 3 / 3

区切ったあとの列のデータ形式を選択してください。

列のデータ形式

☒ G/標準(G)
☐ 文字列(I)
☐ 日付(D): YMD
☐ 削除する(I)

[G/標準] を選択すると、数字は数値に、日付は日付形式の値に、その他の値は文字列に変換されます。

詳細(A)...

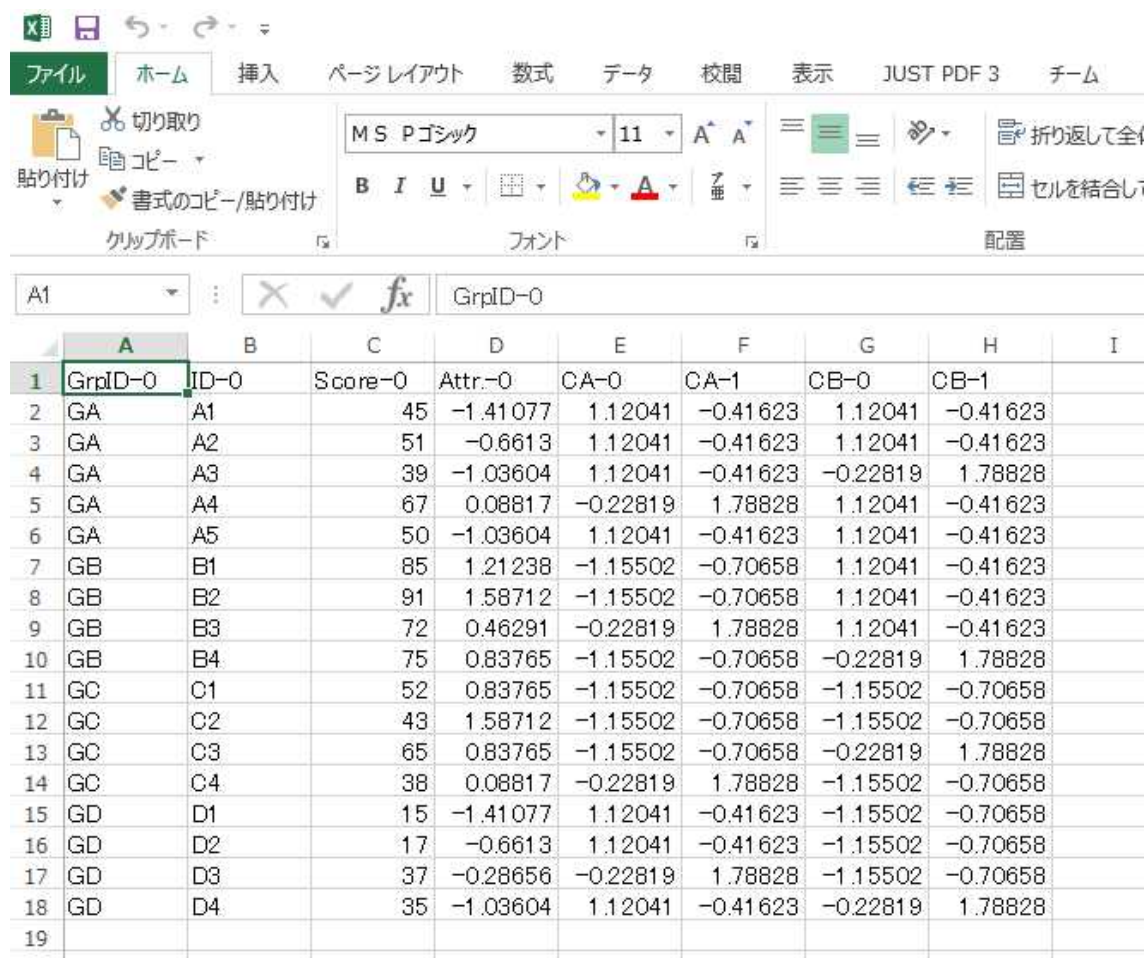
データのプレビュー(P)

G/標準	G/標準	G/標準	G/標準	G/標準	G/標準	G/標準	G/標準
GrpID-0	ID-0	Score-0	Attr.-0	CA-0	CA-1	CB-0	CB-1
GA	A1	45	-1.41077	1.12041	-0.41623	1.12041	-0.41623
GA	A2	51	-0.66130	1.12041	-0.41623	1.12041	-0.41623
GA	A3	39	-1.03604	1.12041	-0.41623	-0.22819	1.78828
GA	A4	67	0.08817	-0.22819	1.78828	1.12041	-0.41623

キャンセル < 戻る(B) 次へ(N) > 完了(E)

図 6

表示されたページ（図6）において、データの表示などに問題がないことを確認して、「完了」をクリックする。図7のように Excel にデータが取り込まれて表示される。



	A	B	C	D	E	F	G	H	I
1	GrpID-0	ID-0	Score-0	Attr.-0	CA-0	CA-1	CB-0	CB-1	
2	GA	A1	45	-1.41077	1.12041	-0.41623	1.12041	-0.41623	
3	GA	A2	51	-0.6613	1.12041	-0.41623	1.12041	-0.41623	
4	GA	A3	39	-1.03604	1.12041	-0.41623	-0.22819	1.78828	
5	GA	A4	67	0.08817	-0.22819	1.78828	1.12041	-0.41623	
6	GA	A5	50	-1.03604	1.12041	-0.41623	1.12041	-0.41623	
7	GB	B1	85	1.21238	-1.15502	-0.70658	1.12041	-0.41623	
8	GB	B2	91	1.58712	-1.15502	-0.70658	1.12041	-0.41623	
9	GB	B3	72	0.46291	-0.22819	1.78828	1.12041	-0.41623	
10	GB	B4	75	0.83765	-1.15502	-0.70658	-0.22819	1.78828	
11	GC	C1	52	0.83765	-1.15502	-0.70658	-1.15502	-0.70658	
12	GC	C2	43	1.58712	-1.15502	-0.70658	-1.15502	-0.70658	
13	GC	C3	65	0.83765	-1.15502	-0.70658	-0.22819	1.78828	
14	GC	C4	38	0.08817	-0.22819	1.78828	-1.15502	-0.70658	
15	GD	D1	15	-1.41077	1.12041	-0.41623	-1.15502	-0.70658	
16	GD	D2	17	-0.6613	1.12041	-0.41623	-1.15502	-0.70658	
17	GD	D3	37	-0.28656	-0.22819	1.78828	-1.15502	-0.70658	
18	GD	D4	35	-1.03604	1.12041	-0.41623	-0.22819	1.78828	
19									

図 7

図7のデータから、「GrpID-0」と「Score-0」の列を削除する。例えば、「GrpID-0」の列を削除するときは、まず「A」をクリックして「GrpID-0」の列をマークする（図8）。

	A	B	C	
1	GrpID-0	ID-0	Score-0	A
2	GA	A1	45	
3	GA	A2	51	
4	GA	A3	39	
5	GA	A4	67	
6	GA	A5	50	
7	GB	B1	85	
8	GB	B2	91	
9	GB	B3	72	
10	GB	B4	75	
11	GC	C1	52	
12	GC	C2	43	
13	GC	C3	65	
14	GC	C4	38	
15	GD	D1	15	
16	GD	D2	17	
17	GD	D3	37	
18	GD	D4	35	
19				
20				

図 8

次に「削除 | シートの列を削除」を選んでクリックする（図 9）。



図 9

指定した列が削除されて、図 10 のようになる。

	A	B	C	D	E	F	G	H
1	ID-0	Score-0	Attr.-0	CA-0	CA-1	CB-0	CB-1	
2	A1	45	-1.41077	1.12041	-0.41623	1.12041	-0.41623	
3	A2	51	-0.6613	1.12041	-0.41623	1.12041	-0.41623	
4	A3	39	-1.03604	1.12041	-0.41623	-0.22819	1.78828	
5	A4	67	0.08817	-0.22819	1.78828	1.12041	-0.41623	
6	A5	50	-1.03604	1.12041	-0.41623	1.12041	-0.41623	
7	B1	85	1.21238	-1.15502	-0.70658	1.12041	-0.41623	
8	B2	91	1.58712	-1.15502	-0.70658	1.12041	-0.41623	
9	B3	72	0.46291	-0.22819	1.78828	1.12041	-0.41623	
10	B4	75	0.83765	-1.15502	-0.70658	-0.22819	1.78828	
11	C1	52	0.83765	-1.15502	-0.70658	-1.15502	-0.70658	
12	C2	43	1.58712	-1.15502	-0.70658	-1.15502	-0.70658	

図 1 0

同様に、列「Score-0」を削除する（図 1 1）。

	A	B	C	D	E	F	G
1	ID-0	Attr.-0	CA-0	CA-1	CB-0	CB-1	
2	A1	-1.41077	1.12041	-0.41623	1.12041	-0.41623	
3	A2	-0.6613	1.12041	-0.41623	1.12041	-0.41623	
4	A3	-1.03604	1.12041	-0.41623	-0.22819	1.78828	
5	A4	0.08817	-0.22819	1.78828	1.12041	-0.41623	
6	A5	-1.03604	1.12041	-0.41623	1.12041	-0.41623	
7	B1	1.21238	-1.15502	-0.70658	1.12041	-0.41623	
8	B2	1.58712	-1.15502	-0.70658	1.12041	-0.41623	
9	B3	0.46291	-0.22819	1.78828	1.12041	-0.41623	
10	B4	0.83765	-1.15502	-0.70658	-0.22819	1.78828	
11	C1	0.83765	-1.15502	-0.70658	-1.15502	-0.70658	
12	C2	1.58712	-1.15502	-0.70658	-1.15502	-0.70658	
13	C3	0.83765	-1.15502	-0.70658	-0.22819	1.78828	
14	C4	0.08817	-0.22819	1.78828	-1.15502	-0.70658	
15	D1	-1.41077	1.12041	-0.41623	-1.15502	-0.70658	
16	D2	-0.6613	1.12041	-0.41623	-1.15502	-0.70658	
17	D3	-0.28656	-0.22819	1.78828	-1.15502	-0.70658	
18	D4	-1.03604	1.12041	-0.41623	-0.22819	1.78828	
19							

図 1 1

図 1 1 のデータファイルを、以下の手順で保存する。

まず、「名前を付けて保存」メニューを選ぶ（図 1 2）。



図 1 2

「ファイルの種類」は、「テキスト（タブ区切り）(*.txt)」でよい（図 1 3）。

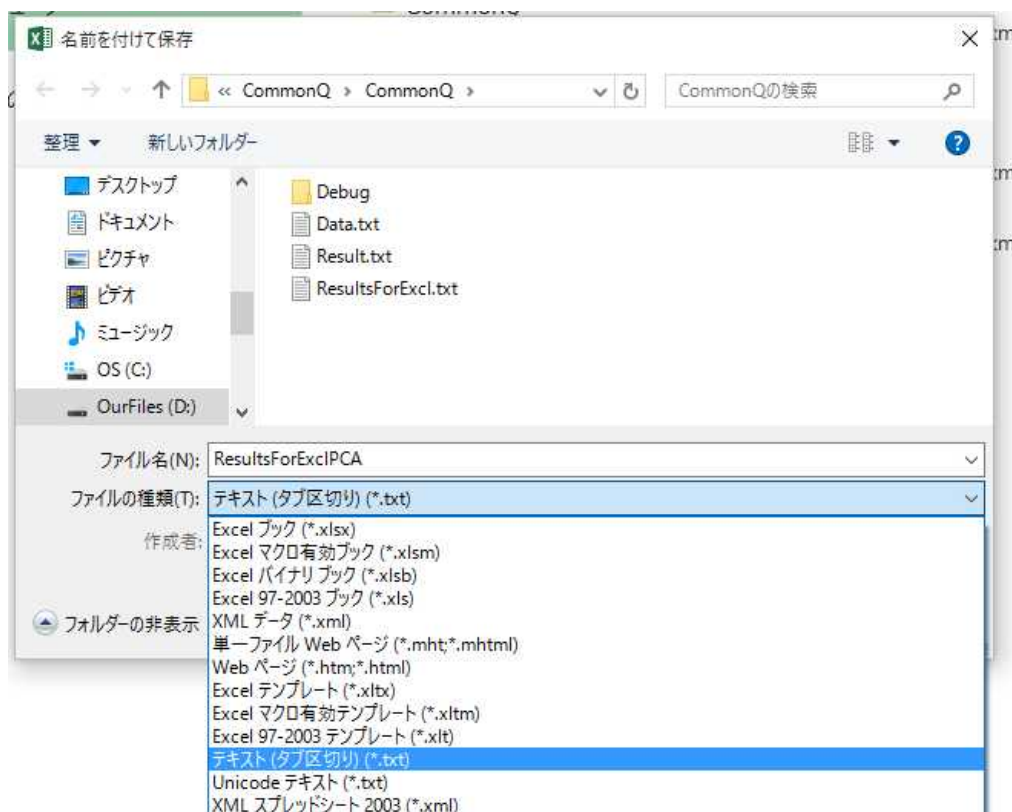


図 1 3

「保存」ボタンをクリックすると、図 1 4 のダイアログフォームが表示されるが、「はい」をクリックする。

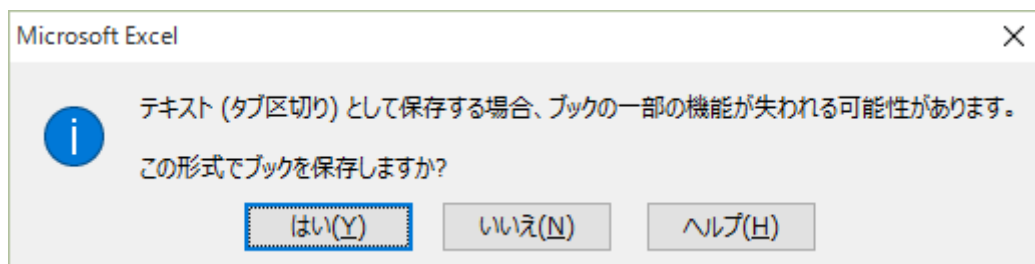


図 1 4

保存したファイルを「メモ帳」で開くと図 1 5 のようになっている。

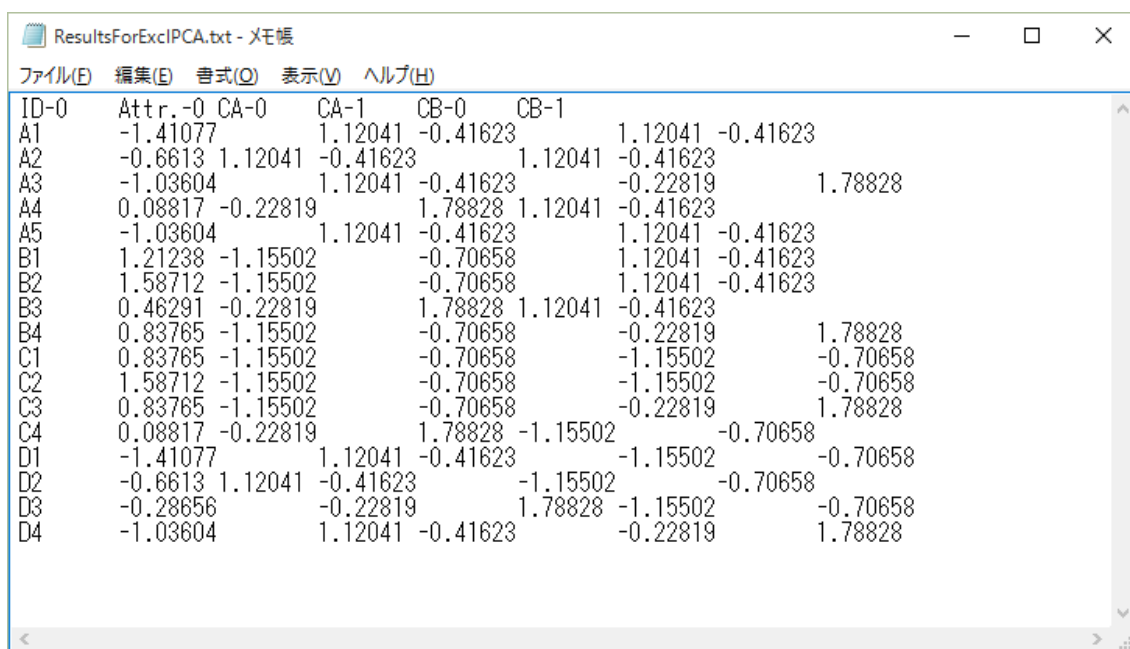


図 1 5

タブ区切りであるので、データの桁数の違いで不揃いになっている。このままでも、ウェブサイトのプログラムの入力ファイルとしては問題ないが、見た目を考えて揃えて、ウェブサイトの主成分分析プログラム用入力ファイルの形式に従って作成すると図 1 6 のようになる。

ResultsForExclPCA.txt - メモ帳

ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)

```
/
    変数の数
6

/
ID-0   Attr.   変数のラベル
      CA-0   CA-1   CB-0   CB-1

/ ID-0   Attr.   CA-0   CA-1   CB-0   CB-1
A1      -1.41077  1.12041 -0.41623  1.12041 -0.41623
A2      -0.6613  1.12041 -0.41623  1.12041 -0.41623
A3      -1.03604  1.12041 -0.41623  -0.22819  1.78828
A4       0.08817  -0.22819  1.78828  1.12041 -0.41623
A5      -1.03604  1.12041 -0.41623  1.12041 -0.41623
B1       1.21238  -1.15502 -0.70658  1.12041 -0.41623
B2       1.58712  -1.15502 -0.70658  1.12041 -0.41623
B3       0.46291  -0.22819  1.78828  1.12041 -0.41623
B4       0.83765  -1.15502 -0.70658  -0.22819  1.78828
C1       0.83765  -1.15502 -0.70658  -1.15502 -0.70658
C2       1.58712  -1.15502 -0.70658  -1.15502 -0.70658
C3       0.83765  -1.15502 -0.70658  -0.22819  1.78828
C4       0.08817  -0.22819  1.78828  -1.15502 -0.70658
D1      -1.41077  1.12041 -0.41623  -1.15502 -0.70658
D2      -0.6613  1.12041 -0.41623  -1.15502 -0.70658
D3      -0.28656  -0.22819  1.78828  -1.15502 -0.70658
D4      -1.03604  1.12041 -0.41623  -0.22819  1.78828

/
```

図 1 6