

## 【補 1】 Excel の操作

以下では、Excel シートに直接関数を書き込むことはしないで、データ行列全体を選択し、それを対象にして、プログラムを使って、様々な変換や分析を行います。プログラムは Excel の VBA (Visual Basic for Application) を使います。

私たちが開発してきた NUMEROS のすべてのプログラムコードを解説することはスペースの関係でできません。また、プログラムは随時改訂しているため、バージョンがすぐ古くなってしまいます。一定のコーディングを習得した後では、全体のプログラムの解説は不要になると思います。そこで、プログラムでは NUMEROS の主要部だけを取り上げて簡単に説明します。ここでは、行列を受け取って一定の演算をした行列を返す一連の関数を「行列関数」とよぶことにします。行列関数の引数は Variant 型の変数です。

このテキストでは、おもに言語の変化・変異を数量的に分析する方法を説明しますので、プログラムで用いる様々な関数やコードの規則については、自分に適した難易度の本を選択して、各自自習しておいてください。Excel の Help やウェブで提供されている多くの情報も役立ちます。

### 9.1. 相関係数

(1) 次のデータを使用します（前節と同じものです）。

鍵語	1 Madrid	2 Sevilla
a	151	163
con	38	45
de	202	195
en	105	81
por	54	45

(2) 次の計算をします。

- B7 =SIM(B2:B6)
- B8 =AVERAGE(B2:B6)
- B9 =STDEVP(B2:B6)

(3) B7:B9 をコピーし， C7 に貼付けます。

	A	B	C
1	鍵語	1 Madrid	2 Sevilla
2	a	151	163
3	con	38	45
4	de	202	195
5	en	105	81
6	por	54	45
7	SUM	550	529
8	MEAN	110.000	105.800
9	SD	60.811	62.027

(4) D2 に標準得点の式を入れます。

$$D2 = (B2 - B\$8) / B\$9$$

(5) D2 を D2:E6 にコピー。桁数が不統一だと比較しにくいので D, E 列の書式を小数点以下 3 とします。

	A	B	C	D	E
1	鍵語	1 Madrid	2 Sevilla	v-1(sm)	v-2(sm)
2	a	151	163	0.674	0.922
3	con	38	45	-1.184	-0.980
4	de	202	195	1.513	1.438
5	en	105	81	-0.082	-0.400
6	por	54	45	-0.921	-0.980

(6) B7:C9 をコピーして D7 に貼付けます。

	A	B	C	D	E
1	鍵語	1 Madrid	2 Sevilla	v-1(sm)	v-2(sm)
2	a	151	163	0.674	0.922
3	con	38	45	-1.184	-0.980
4	de	202	195	1.513	1.438
5	en	105	81	-0.082	-0.400
6	por	54	45	-0.921	-0.980
7	SUM	550	529	0	0
8	MEAN	110.000	105.800	0.000	0.000
9	SD	60.811	62.027	1.000	1.000

これで正しく標準化されたことがわかります。次に，これらの数値をもとに相関係数を求めてみましょう。まず，それぞれの項目の標準得点の積と全体の積平均を求めます。

F2		fx =D2*E2				
	A	B	C	D	E	F
1	鍵語	1 Madrid	2 Sevilla	v-1(sm)	v-2(sm)	積和
2	a	151	163	0.674	0.922	0.622
3	con	38	45	-1.184	-0.980	1.161
4	de	202	195	1.513	1.438	2.176
5	en	105	81	-0.082	-0.400	0.033
6	por	54	45	-0.921	-0.980	0.903
7	SUM	550	529	0	0	4.894
8	MEAN	110.000	105.800	0.000	0.000	0.979

F2 = D2\*E2

F2 を(F3:F6)にコピー

(B7:B8)を(F7:F8)にコピー

これで標準得点をもとに相関係数を求めることができました。

結果を確認するために、Excel 関数を使って相関係数を算出し比較してみましょう。Excel には CORREL という関数が用意されており、対象となる2つのデータをコンマ区切りで選択します。

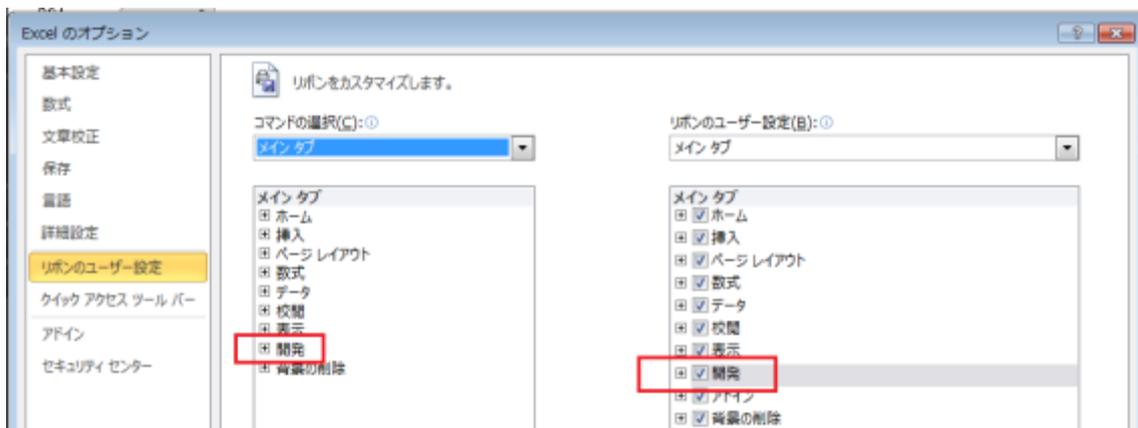
B10 =CORREL(B2:B6,C2:C6)

B10		fx =CORREL(B2:B6,C2:C6)				
	A	B	C	D	E	F
1	鍵語	1 Madrid	2 Sevilla	v-1(sm)	v-2(sm)	積和
2	a	151	163	0.674	0.922	0.622
3	con	38	45	-1.184	-0.980	1.161
4	de	202	195	1.513	1.438	2.176
5	en	105	81	-0.082	-0.400	0.033
6	por	54	45	-0.921	-0.980	0.903
7	SUM	550	529	0	0	4.894
8	MEAN	110.000	105.800	0.000	0.000	0.979
9	SD	60.811	62.027	1.000	1.000	
10	Correl	0.979				

F8 と B10 の値が同じになることを確認しましょう。

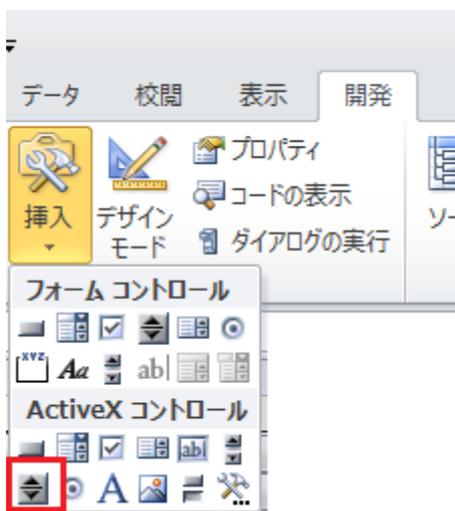
それぞれの特徴を見るために値を操作するのに「スピンボタン」を使うと便利です。

(1) はじめにリボンに「開発」タブを設定します。◆「ファイル」→「オプション」→「リボンのユーザー設定」→]を選択し、「リボンのユーザー設定」で「メインタブ」の「開発」のチェックボックスをオンにします。

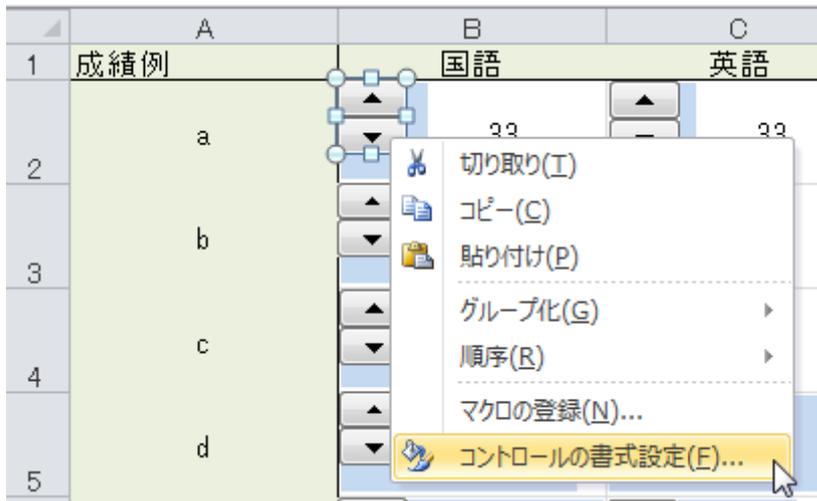


Excel 2007 : 「Office ボタン」 → 「Excel のオプション」 → 「基本設定」 → 「[開発]タブをリボンに表示する」をチェック

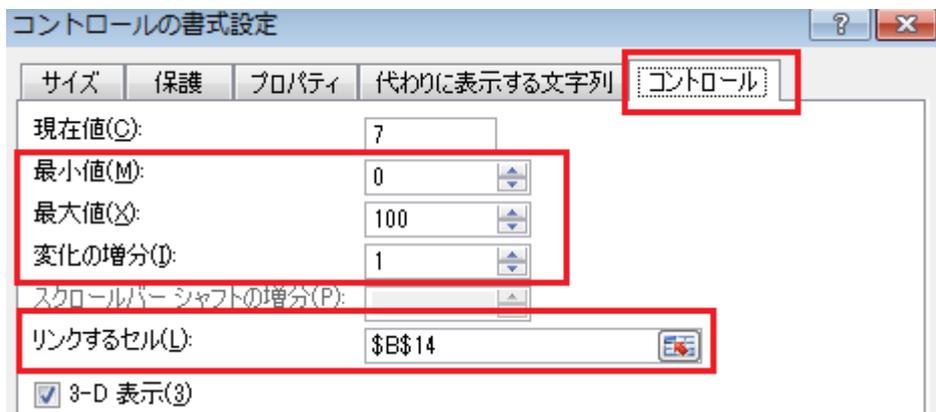
(2) 「開発」 → 「コントロール」 → 「挿入」 → 「フォームコントロール」の中のスピノボタンをクリック → シート内の適当な位置にドラッグして配置します。



(3) シートに配置したスピノボタンを右クリック → 「コントロールの書式設定」



(4) 「コントロール」タブ→「最小値」「最大値」「変化の増分」「リンクするセル」を設定します。「リンクするセル」にスピンボタンによる入力の結果が表示されます。



(5) スピンボタンなどのコントロールは右クリックすることにより、大きさの変更、ドラッグ、コピー、などが可能になります。

スピンボタンは便利なのですが、たとえば 1 から 100 まで移動するときは大変です。スピンボタンをつけたらそれではか値が操作できなくなるというわけではなく、直接セルに 100 と記入することもできます。

## 9.2. 連関係数

連関係数を使ってデータを比較するにはまず量的なデータを質的なデータに変換する必要があります。これには IF 関数を使えば便利です。例として次のデータを使用します。

	A	B	C	D	E	F	G	H	I	J	K
1	語	基準値	同値		基準値	同値		a++	b+	c+	d--
2	abajo	5	10		1	1		1	0	0	0
3	abandonar	9	6		1	1		1	0	0	0
4	abandono	0	0		0	0		0	0	0	1
5	abarcar	1	0		1	0		0	1	0	0
6	abastecimiento	2	0		1	0		0	1	0	0
7	abatir	0	1		0	1		0	0	1	0
8	abeja	2	3		1	1		1	0	0	0
9	abertura	0	0		0	0		0	0	0	1
10	abismo	0	0		0	0		0	0	0	1
11	abnegación	0	0		0	0		0	0	0	1
12	abogado	3	6		1	1		1	0	0	0
13	abonar	5	0		1	0		0	1	0	0
14	abono	0	0		0	0		0	0	0	1
15	abordar	0	0		0	0		0	0	0	1
16	aborrecer	0	6		0	1		0	0	1	0
17				基準値	0	和		4	3	2	6

(1) はじめに、量的データの質化の基準を設定します。

A17 を質的データに変換するための基準値とします。この値よりも大きい場合、「1」に変換するというルールにします。0 よりも大きいときに変換する場合は F17=0 と記入しておきます。

(2) IF 関数を使って量的データ(B2)を質的データ(E2)に変換します。

$$E2=IF(B2>F17, 1, 0)^1$$

この式の意味は、E2 が基準値の値(0)よりも大きい場合は、1 をそれ以外は 0 を返す、ということです。

(3) E2 を E2:F16 にコピーします。これで 0 より大きい値を 1 と表示することができました。

次に、さきほどの変換の結果を基に、共通して使われているもの、一方だけ使われているもの、どちらも使われていないものを集計しましょう。

<sup>1</sup> ここでは基準値を動かすことができるように\$F\$17 としましたが、下記のように\$F\$17 を使用しなくても同じ結果を得ることができます。

$$E2=IF(B2>0, 1, 0)$$

AVERAGE    x ✓ f    =IF(AND(\$E2=1, \$F2=1), 1, 0)

	A	B	C	D	E	F	G	H	I	J	K
1	語	発音	意味		発音	意味		a++	b+-	c+-	d--
2	abajo	5	10		1	1		0	1	1	1
3	abandonar	9	6		1	1		1	0	0	0
4	abandono	0	0		0	0		0	0	0	1
5	abarcar	1	0		1	0		0	1	0	0
6	abastecimiento	2	0		1	0		0	1	0	0
7	abatir	0	1		0	1		0	0	1	0
8	abeja	2	3		1	1		1	0	0	0
9	abertura	0	0		0	0		0	0	0	1
10	abismo	0	0		0	0		0	0	0	1
11	abnegación	0	0		0	0		0	0	0	1
12	abogado	3	6		1	1		1	0	0	0
13	abonar	5	0		1	0		0	1	0	0
14	abono	0	0		0	0		0	0	0	1
15	abordar	0	0		0	0		0	0	0	1
16	aborrecer	0	6		0	1		0	0	1	0

(5) はじめに E2 と F2 を対象としてデータを入力します。

H2 =IF(AND(\$E2=1, \$F2=1), 1, 0)

この式の意味は、E2（手紙）と F2（演劇）が共に 1 の場合、1 を返し、それ以外は 0 にする、ということです。AND を使って複数の条件を指定していることに注意してください。

H2 を I2:K2 にコピーして、一部を次のように修正します。

I2=IF(AND(\$E2=1, \$F2=0), 1, 0)

J2=IF(AND(\$E2=0, \$F2=1), 1, 0)

K2 IF(AND(\$E2=0, \$F2=0), 1, 0)

(6) H2:K2 を H2:K16 にコピーします。

(7) G17 を書き込み、SIM で H17:K17 を計算します。

H17 =SIM(H2:H16)

H17 を I17:K17 にコピーします。I17 =SIM(I2:I16)

J17 =SIM(J2:J16)

K17 =SIM(K2:K16)

最終的には次のような値になります。

	a++	b+-	c+-	d--
和	4	3	2	6

これで四象限での集計が完了です。

## 各種の連関係数

それでは各種の連関係数を計算してみましょう。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	語	出現	重複	出現	重複	a++	b+-	c-+	d--				Pearson = Phi	0.577	r
2	abajo	5	10		1	1		1	0	0	0				
3	abandonar	9	6		1	1		1	0	0	0				
4	abandono	0	0		0	0		0	0	0	1				
5	abarcar	1	0		0	0		0	0	0	1				
6	abastecimiento	2	0		1	0		0	1	0	0				
7	abrir	0	1		0	0		0	0	0	1				
8	abeja	2	3		1	1		1	0	0	0				
9	abertura	0	0		0	0		0	0	0	1				
10	abismo	0	0		0	0		0	0	0	1				
11	abnegación	0	0		0	0		0	0	0	1				
12	abogado	3	6		1	1		1	0	0	0				
13	abonar	5	0		1	0		0	1	0	0				
14	abono	0	0		0	0		0	0	0	1				
15	abordar	0	0		0	0		0	0	0	1				
16	aborrer	0	6		0	1		0	0	1	0				
17					基準値	1	和	4	2	1	8		類似係数		
18															
19													Co-occurrence	4.000	a
20													Simple match	0.800	(a+d)/(a+b+c+d)
21													Russel and Rao	0.267	a/(a+b+c+d)
22													Jaccard	0.571	a/(a+b+c)
23													Dice	0.727	2a/(2a+b+c)
24													Yule	0.882	(ad-bc)/(ad+bc)
25													Hamann	0.600	((a+d)-(b+c))/((a+d)+(b+c))
26													Phi	0.577	(ad-bc)/sqrt((a+b)*(a+c)*(d+b)*(d+c))
27													Ochiai	0.730	a/sqrt((a+b)*(a+c))
28													Preference	0.455	(2a - (b+c))/(2a+(b+c))
29															
30															
31															

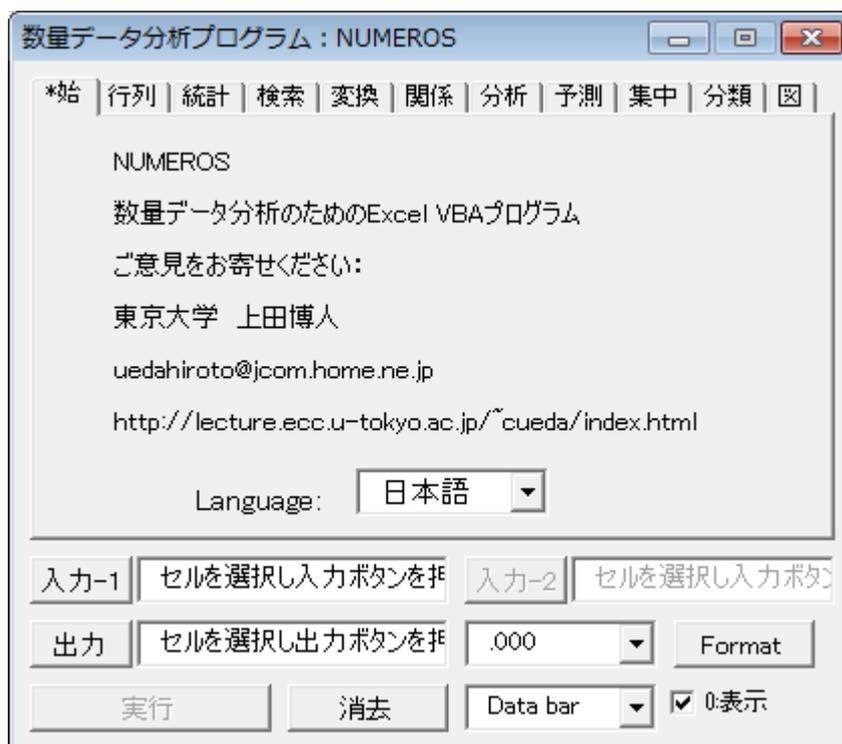
M19:M27 でそれぞれの係数を求めます。

- (1) 共起回数 : M19=H17
- (2) Simple match 係数 : M20=(H17+K17)/(H17+I17+J17+K17)
- (3) Russel and Rao 係数 : M21=H17/(H17+I17+J17+K17)
- (4) Jaccard 係数 : M22=H17/(H17+I17+J17)
- (5) Yule 係数 : M23=((H17\*K17)-(I17\*J17))/((H17\*K17)+(I17\*J17))
- (6) Hamann 係数 : M24=((H17+K17)-(I17+J17))/((H17+K17)+(I17+J17))
- (7) Phi 係数 :  

$$M25 = \frac{(H17 * K17) - (I17 * J17)}{\sqrt{((H17 + I17) * (H17 + J17) * (I17 + K17) * (J17 + K17))}}$$
- (8) Ochiai 係数 : M26=H17/SQRT((H17+I17)\*(H17+J17))
- (9) Prominence 係数 : M27=(H17/(H17+I17)+H17/(H17+J17))/2
- (10) Preference 係数 : M30=(2\*H17-I17-J17)/(2\*H17+I17+J17)

## 9.3. マクロプログラム

Excel VBA によるマクロプログラム(NIMEROS.xlsm)を使って言語資料の分析をします。そのときに用いる簡単な統計の概念についても説明します。ここで扱う統計的手法の多くは一般の参考書(→参考書)で説明されているものですが、一部は私たちが提案する手法や計算法も含めます。



Excel ファイル NUMEROS.xlsm は Excel2010, Excel2013 で動作します。起動時にマクロを有効にしてください。

すべての演算の対象は次のような行列です。シート内の一定の範囲に展開されているデータは条件として次のような構成にしなければなりません。

A	v1	v2	v3	v4	v5
d1	10	19	14	7	12
d2	11	7	10	0	1
d3	0	0	1	12	1
d4	0	1	2	3	3

この行列は「タイトルセル」(A), 「タイトル行」(v1, v2, ..., v5), 「タイトル列」(d1, d2, d3, d4), 「データ」(10, 19, ..., 3, 3)という構成になります。この行列を「データ行列」 $D_{np}$ とよびます。 $D_{np}(1,1)=10$ ,  $D_{np}(1,2)=19$ , という行列です。タイトルセルは  $D_{np}(0,0)$  に, タイトル行は  $D_{np}(0,1)$ ,  $D_{np}(0,2)$ , ... に, タイトル列は  $D_{np}(1,0)$ ,  $D_{np}(2,0)$ , ... に格納されます。

シート内の上のようなデータの一部のセルを選択し, NUMEROS の「入力 1」ボタンを押すと, その行列全体が選択されます。その後, この行列から離れたセルをクリックし, 「出力」ボタンを押すと, 選択されたセルが出力の開始セルになります。ここが出力領域の上・左端になります。このテキストで説明したさまざまな数値データ処理のメニューはタブを開いて, 選択してください。「始」以外のタブを開くと, 「実行」ボタンがクリックできる状態になります。

次が「行列入力」のコードの主要部です。これを適当なコマンドボタンと連結させます。

```
Sub MATRIX_INPIT() '● 行列入力()
  I_np = Selection: N = Ir(I_np) - 1: P = Ic(I_np) - 1
  ReDim G_np(N, P)
  For i = 0 To N: For j = 0 To P
    D_np(i, j) = I_np(i + 1, j + 1) '入力行列
  Next: Next
  I_np=D_np
End Sub
```

```
Function Ir(F_np): Ir = IBound(F_np, 1): End Function '行数
Function Ic(F_np): Ic = IBound(F_np, 2): End Function '列数
```

バリエーション型変数  $I_{np}$  に選択範囲(Selection)の内容を代入し、 $N$  を行数、 $P$  を列数とします。 $Ir$  と  $Ic$  はそれぞれ行列の行数と列数を返すユーザー定義関数です(Function ... End Function)。 $Ibound(*, *)$ は Excel 関数です。概説書やウェブで確認してください。 $I_{np}$  にはシートに選択された内容が、そのまま格納されてるので、 $I_{np}(1,1)=""O.S.""$ になります。行についても列についても、番地を1つずつ減らして移動したものが  $D_{np}$  です。For ... Next を0から  $N$ 、0 から  $P$  まで繰り返して移動します。その結果  $D_{np}(0,0)=""O.S.""$ 、 $D_{np}(1,1)=10$ 、...、 $D_{np}(4,5)=3$  になります。この  $D_{np}$  を改めて  $I_{np}$  に代入します。この  $I_{np}$  が新たな入力行列になります。これに、以下で説明する一定の処理をして、できあがった行列を次の「行列出力」に渡します。ユーザーは、適当なコマンドボタンと連結した次の「出力位置」を実行しておき、出力位置の情報を確保しておきます。

```
Sub OITPIT_POSITION() '● 出力位置
  OpSheet% = ActiveSheet.Index '出力シート番号
  OpPosition$ = ActiveCell.Address(bF, bF) '出力位置のアドレス
  OpRowNum& = ActiveCell.Row '出力位置の行番号
  OpColNum& = ActiveCell.Column '出力位置の列番号
End Sub
```

これらを、「入力」「出力」ボタンに連動させます。

次が行列出力の主要部です。

```
Sub MATRIX_OITPIT(F_np, F1&, F2&) '● 行列出力
  Sheets(OpSheet%).Select '出力シートを選択
  Cells(OpRowNum&, OpColNum&).Offset(F1, F2).Select '出力セルを選択
```

```

Selection.Resize(Ir(Fnp) + 1, Ic(Fnp) + 1).Select '拡大リサイズ
Selection = Fnp '行列を代入
End Sub

```

「出力位置」で得た出力シート番号で出力シートを選択し、出力位置のセルを選択し、F1, F2 で位置を移動し(複数の出力がある処理のためです)、さらに、行列の行数と列数にそれぞれ 1 を加えた数だけ拡大リサイズした選択位置に行列 F<sub>np</sub> を代入します。これで指定されたシート位置に行列が出力されます。

これを、たとえば、次のように出力部から呼び出します。

```

Sub MATRIX()
Dim Xnp, Ynp: Xnp = Inp: Ynp = Dnp
Select Case fN.lstMat.ListIndex + 1 'ListBox
Case 1: Onp = Iv(fN.cboScalar1)      '0.単位ベクトル_Iv
Case 2: Onp = Im(fN.cboScalar1)     '1.単位行列_Im
...
End Select
Call MATRIX_OITPIT(Onp, 0, 0)
End Sub

```

ExcelVBA のプログラミングについては多くの書籍が出版されています。また、ウェブサイトでも情報がたくさん得られます。以下では、NIMEROS.xlsm の主要部のコードについて解説します。コード全体については VBA Editor を開いてください。

## 【補 2】 参考書

### 基礎

- 池田央(1976)『統計的方法 I 基礎』新曜社 (記述統計について数理と具体例で納得できる説明がなされています)
- 石村貞夫(1995)『グラフ統計のはなし』東京図書 (やさしく説明してあるのでクラスター分析がどのようなものかがわかります)
- 芝祐順・渡部洋・石塚智一(1984)『統計用語辞典』新曜社(一般的な統計学の用語と英語を知るのに便利です。参考書目も充実しています)
- 清水誠(1996)『データ分析はじめの一步』講談社ブルーバックス (書名のとおりデータ分析法を学ぶために最初に読むとよいでしょう)
- 東京大学教養学部統計学教室(1991)『統計学入門』(東京大学出版会)(理論をしっかりと理解するために役立つ入門書です)
- 渡辺美智子・神田智弘(2008)『統計データ分析』秀和システム (Excelを使った記述統計学の分析法を具体的に説明しています)

### 応用・開発

- 足立堅一(2005)『多変数解析入門：線形代数から多変数解析へ』篠原出版新社。(多変数解析に向けて線形代数の基本を語りかけるような文面で丁寧に説明されています)
- Anderberg, Michael R. (1973) *Cluster analysis for applications*. New York, Academic Press. 西田英朗・佐藤嗣二他訳『クラスター分析とその応用』内田老鶴圃(1988). (クラスター分析を知るための古典的な図書です)
- Ávila, R. Samper, J. A. y Ueda, H. (2003) *Pautas y pistas en el análisis del léxico hispanoamericano*. Iberoamericana Vervuert, 278pp. (言語統計分析の方法をスペイン語の語彙バリエーションの研究に応用しました)
- Bertin, Jacques. (1977) *La graphique et le traitement graphique de l'information*. Paris: Flammarion. 森田喬訳『図の記号学』平凡社, 1982. (ここで取り上げた集中分析に関連したことを手作業で実行しています)
- Cahuzac, Philippe. (1980) "La División del español de América en zonas dialectales: Solución etnolingüística o semántico-dialectal." *Lingüística Española Actual*, 10, pp. 385-461. (集中分析で資料を引用しました)
- Hartigan, J. A. (1975) *Clustering Algorithms*. New York. John Wiley & Sons.
- 市原清志 (1990)『バイオサイエンスの統計学』南江堂 (実験を使った検定の方法の説明が具体的でわかりやすく参考になります)
- 長谷川勝也(2001)『はじめての行列とベクトル』技術評論社(高校で数学 III を履修していない人ははじめてに読んでおくとよいでしょう)
- Hahsler, M. / Grun. B. / Hornik, K. / Buchta, Ch. "Introduction to arules – A

computational environment for mining association rules and frequent item sets" <https://lyle.smu.edu/IDA/arules/> (24/2/2016) (連関分析 Association analysis の R のためのパッケージを解説)

- Hoel, P. G. *Introduction to mathematical statistics*. 浅井晃・村上正康訳(1978)『入門数理統計学』培風館 (少し難易度が高いですが、数式だけでなく言葉による説明が多い本です)
- Horst, Paul. (1965) *Factor Analysis of Data Matrices*. Holt, Rinehart and Winston. 柏木繁男・芝祐順・池田央・柳井晴夫訳『コンピュータによる因子分析法』科学技術出版社, 1978. (難解)
- 石川慎一郎(2008)『英語コーパスと言語教育』大修館書店 (第Ⅲ章「コーパスと言語の計量」を共起係数の参考にしました)
- 稲垣宣生(2003)『数理統計学』裳華房 (本格的な統計学の理論書。かなり難解ですが問題解決のヒントが得られます)
- 一石賢(2004)『道具としての統計解析』日本実業出版社 (各種の確率分布関数の難解な数学的導出を独特の語り口で丁寧に説明し「なぜ」という疑問に答えています)
- 小寺平治(2002)『ゼロから学ぶ統計解析』講談社 (書名のとおり統計解析をはじめて勉強する人に適した本。抽象的な確率を具体的な例で説明し、証明もわかりやすい)
- 倉田博史・星野崇宏(2009)『入門統計解析』新生社 (多くの統計数理の証明が丁寧です。ジニ係数が詳しく説明されています)
- 井上勝雄(1998)『パソコンで学ぶ多変数解析の考え方』筑波出版会 (説明がやさしくわかりやすいです。このテキストの数量化Ⅲ類・対応分析のプログラミングで参考にしました)
- 井上勝雄・広川美津雄(2000)『エクセルで学ぶ多変数解析の作り方』筑波出版会 (Excel VBA のプログラミング法の具体的な説明があります)
- 石井俊全(2014)『意味がわかる多変数解析』ベレ出版 (数理の証明がとてもわかりやすく納得できます)
- 岩崎学・吉田清隆(2006)『統計的データ解析入門 線形代数』東京図書 (線形代数のエッセンスを簡潔に説明しています)
- 河口至商 (1978)『多変数解析入門 I, II』森北出版(連関係数行列の説明が参考になります)
- 小林竜一(1967)『社会科学のための数学概説』共立出版(私が1年生のときの数学の教科書。今でも参照しています)
- 三野大來(2001)『統計解析のための線形代数』共立出版 (多変数解析にとって重要な行列演算が簡潔にわかりやすく説明されていて、行列関数によるプログラミングにとっても役立ちます。このテキストでも各所で応用しました)
- 長沢伸也・中山厚穂(2009)『Excel ソルバー多変量解析：因果関係分析・予測手法編』日科技連 (ロジスティック回帰分析の方法を参照しました)

- 永田靖(2005)『統計学のための数学入門 30 講』朝倉書店（文系の高校数学を超える数理についてはこの本を参考にするとよいでしょう）
- 縄田和満(1999)『Excel による線形代数入門』朝倉書店（Excel の操作を通して数理と具体的な方法が学べます）
- 荻野綱男(1980)「敬語における丁寧さの数量化：札幌における敬語調査から(2)」『国語学』vol. 120, pp. 13-24.（クロス集計表の簡易的な両軸の数量化法が提案されています。この数量化された軸を使って集中分析をすることができます）
- 奥野忠一・久米均・芳賀敏郎・吉澤正(1971)『多変量解析法』日科技連（本格的な学術書）
- 奥村晴彦(1986)『パソコンによるデータ解析入門．数理とプログラミング実習』技術評論社．（BASIC 言語によるプログラムがあります。クラスタリングとデンドログラムのマクロはこのプログラムを参考にしました）
- 尾崎隆『ビジネスに活かすデータマイニング』技術評論社(2014)（連関分析についてわかりやすく解説しています。）
- Rosemburg, Ch. H. (1989) *Cluster analysis for researchers*. Robert E. Krieger Publishing Company, Inc. Malabar, Florida. 西田英朗・佐藤嗣二訳『実例クラスター分析』内田老鶴圃(1992).
- 芝祐順(1975)『行動科学における相関分析法』東京大学出版会．（さまざまな相関分析法が簡潔に説明されています。巻末の FORTRAN プログラムが参考になります）
- 芝祐順・南風原朝和(1990)『行動科学における統計解析法』東京大学出版会．（統計の各種指標の理論的な背景がわかります）
- 白井豊(2009)『Excel と VBA による実用数値解析入門』ゆたか創造舎．（固有値と重回帰分析のプログラムを参考にしました）
- 高村大也(2010)『言語処理のための機械学習入門』コロナ社．（数式の意味が例題の具体例でわかるように工夫されています）
- 高橋信(2005)『Excel で学ぶコレスポンデンス分析』オーム社．（対応分析の手順が具体的にわかりやすく説明されています。この本を参照しながらこのテキストのプログラムを作成しました）
- 竹内啓・柳井晴夫(1972)『多変数解析の基礎』東洋経済新報社（Horst ではわかりにくい Varimax 法の理論を詳しく論じています）
- 内田治(2002)『すぐわかる EXCEL による回帰分析』東京書籍（数値例を使って Excel 関数による分析をしているので方法が具体的に理解できません。ロジスティック回帰分析について参照しました）
- 安田三郎・海野道朗(1977)『社会統計学』（改訂 2 版）丸善（クラスター

分析がわかりやすく具体的に解説されています)

\*はじめて行列・ベクトルについて勉強する人は、長谷川(2001)→縄田(1999)→三野(2001)→芝(1975)→白井(2009)の順で読み進めるとよいでしょう。

\*はじめて確率について勉強する人は、小寺(2002)→倉田・星野(2009)→Hoel(1978)→稲垣(2003)の順で読み進めるとよいでしょう。

[FIN]