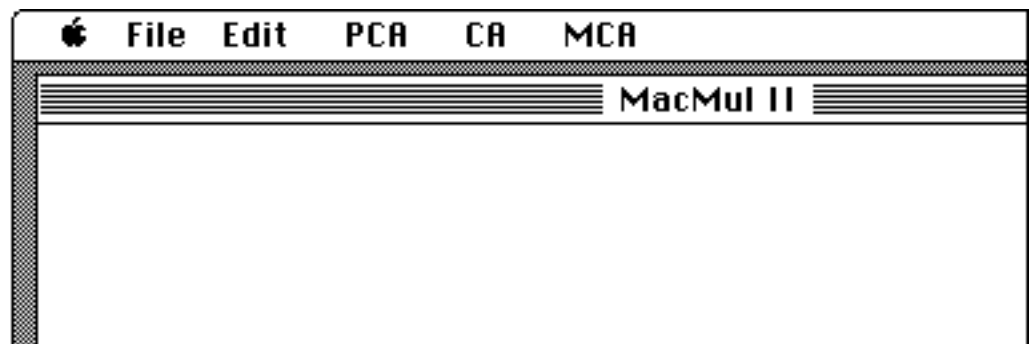


# MacMul

## Multivariate Data Analysis

Principal component analysis  
Correspondence analysis  
Multiple correspondence analysis



**Jean Thioulouse**

U.R.A. CNRS 243 "Biométrie, Génétique et Biologie des populations"  
Université Lyon 1 - I.A.S.B.S.E.- 69622 Villeurbanne Cedex - France  
e-mail : [jean@biomac.univ-lyon1.fr](mailto:jean@biomac.univ-lyon1.fr) - [THIOULOU@FRCISM51.BITNET](mailto:THIOULOU@FRCISM51.BITNET)

3 June 1991

**MACMUL  
MULTIVARIATE ANALYSIS SOFTWARE  
ON THE MACINTOSH**

**Jean THIOULOUSE**

Laboratoire de Biométrie, Génétique et Biologie des populations  
U.R.A. CNRS 243 - Université Lyon 1  
69622 Villeurbanne Cedex - France  
Fax : 72 44 84 66 - Telex : 330722 F  
e-mail : jean@biomac.univ-lyon1.fr - THIOULOU@FRCISM51.BITNET

This document is the MacMul program user manual. It is in six parts: the first is a general introduction, the second describes each of the MacMul "File" menu commands and the following three concern the three menus PCA, CA and MCA. The last contains examples of use.

MacMul must be used together with **GraphMu** (graphics program for the analysis of multivariate data), also available from the author.

Publication of results obtained with this software in scientific reviews is submitted to citation of one of the following papers:

Thioulouse J. (1989). Statistical analysis and graphical display of multivariate data on the Macintosh. *Computer Applications in the Biosciences*, **5**, 4, 287-292.

Thioulouse J. (1990). MacMul and GraphMu : two Macintosh programs for the display and analysis of multivariate data. *Computers and Geosciences*, **16**, 8, 1235-1240.

# Contents

<b>1 - Introduction .....</b>	<b>1</b>
General comments.....	1
Data files .....	1
Files created during an analysis .....	2
<b>2 - Use .....</b>	<b>3</b>
□ Menu .....	3
File menu.....	4
Help .....	5
Transfo. TEXT -> BIN .....	5
Transfo. BIN -> TEXT .....	6
List BIN file.....	6
Print .....	7
Work folder .....	7
Transpose file.....	7
Transfer .....	7
GraphMu .....	8
Quit .....	8
Edit menu .....	8
<b>3 - PCA menu .....</b>	<b>9</b>
Preparation .....	9
Computations .....	11
Row Inertia .....	12
Column inertia .....	14
Reconstitution .....	15
Additional rows .....	16
Additional columns .....	17
PCA on profiles .....	17
<b>4 - CA menu .....</b>	<b>19</b>
Preparation .....	19
Computations .....	20
Row Inertia .....	20
Columns inertia .....	21
Reconstitution.....	21
Additional rows .....	22
Additional columns .....	22
Double discrimination .....	22

<b>5 - MCA menu</b> .....	25
Preparation .....	25
Computations .....	26
Row Inertia .....	27
Column inertia .....	27
Correlation ratios .....	28
Additional rows .....	29
Additional columns .....	29
 <b>6 - Examples of use</b> .....	31
PCA .....	31
PCA : analysis on profiles.....	40
CA: "double discrimination" .....	41
MCA: "Correlation ratios" .....	46

# 1 - INTRODUCTION

## General comments

MacMul is a program for the Macintosh intended to perform the calculations of three basic multivariate analysis methods. These three methods are: principal components analysis (PCA), correspondence analysis (CA) and multiple correspondence analysis (MCA). This user manual supposes that you are already familiar with these methods and that you have some mastery of the Macintosh user interface (scrolling menus, dialog windows, mouse, etc.). If this is not the case, before using MacMul you should read some basic texts in these two fields.

The principle behind the writing of the MacMul program was to make the carrying out of the three analysis methods as simple as possible and with their input and output as comparable as possible. Moreover, we tried to make the program conform to the general rules respected by most Macintosh programs (standardized menus, dialog windows, etc.). We also tried to give the user maximum freedom in the preliminary stages (handling of data files) and the later stages (editing results and graphic representations). MacMul does not produce graphic representations: the GraphMu program must be used for this.

## Data files

The data files used by MacMul can be created by all Macintosh programs which can create "Text" files. This is true for both spreadsheets (e.g. Excel™, StatView™,...) and word processing (MacWrite™, Word™, Edit™, ...). Figure 1.1 shows the icons of the text files created by some classical programs.



*Figure 1.1:  
Icons of "Text" files created by some Macintosh applications*

In all the data files, the rows must correspond to individuals on which a series of variables (in a column) has been measured.

The maximum number of columns authorized varies as a function of the number of rows. The exact limit can be calculated as follows: if  $L$  is the number of rows,  $C$  the number of columns and  $M$  the smaller of these two values, one must have

$L+4*C+M*M \leq 500,000$ . As an example, for 600 columns the maximum number of rows is 137,600 (because  $137,600 + 4*600 + 600*600 = 500,000$ ). This calculation is the same for the three methods (PCA, CA and MCA). In MCA however, the number of columns one must use is not the number of qualitative variables, but the sum of the number of categories for all the variables.

The number of rows of the data file may be less than the number of columns, for all three methods. The calculations are in all cases performed in the lower dimensional space,  $\min(L,C)$ .

## Files created during an analysis

The files of results created by MacMul (e.g. factorial coordinate files) are a special type of file (called "binary files"); they have an icon which allows them to be recognized (figure 1.2).



Figure 1.2:  
*Icon of binary files created by MacMul.*

These binary files cannot be read by programs other than MacMul, GraphMu and ADECO (ecological data analysis program library in compiled Microsoft QuickBasic, interfaced with HyperCard). They can however be translated into text files, which can then be processed with a spreadsheet or word processor (these files have the same icon as the "Edit" file in figure 1.1).

MacMul automatically creates a certain number of files (binary files and text files), containing the analysis results. All the files are created in the "work folder", which is the folder where one finds the data file used for the analysis. The name of these files is automatically constructed from a string of characters, called "Analysis title", which is given by the user at the beginning of the calculations: the program adds four-letter extensions to this file, allowing it to recall the file contents. These four letters depend on the method used and the file contents. Table 1 shows some examples of file names created by MacMul.

Method:	Row coordinates	Column coordinates
PCA	xxx.CPLI	xxx.CPCO
CA	xxx.FCLI	xxx.FCCO
MCA	xxx.MCLI	xxx.MCCO

Table 1:  
*Examples of file names created by MacMul for the three method. The analysis title here is "xxx" and the files contain the factorial coordinates for rows or columns*

## 2 - USE

For each analysis the user must perform two stages, **preparation** and **computations**. These two stages allow the user to obtain the factorial coordinate files needed to produce the graphics (e.g. factorial planes). It is then possible to perform further complementary calculations, grouped under the name of "*aids to interpretation*". The most classical of these aids are:

- inertia analyses, for rows and columns,
- reconstitution of initial data,
- projection of additional rows and columns.

Other commands are available, depending on the method used.

During the computation stage, MacMul displays the list of eigenvalues obtained during the diagonalisation of the scalar product matrix. The user must then indicate the "number of factors to keep". This number of factors will determine the **number of columns** of the factorial coordinate files.

MacMul comes in two versions: one for 68020/63030 microprocessors and 68881/68882 math co-processors (i.e. for Macintosh II, IIfx, IIfx, IIfx, SE/30, and IIfx *with optional math co-processor*), and another for "small" Macintosh (Plus, SE, Portable, Classic, LC, and IIfx *without math co-processor*). Only the first version exists with an English language interface. It needs 3Mb of main memory, while the second one only needs 800Kb.

First of all we will study the commands of each of the menus proposed by MacMul (figure 2.1): "Apple" menu (□), "File" menu, "Edit" menu and "PCA", "CA" and "MCA" menus.



Figure 2.1:  
MacMul menu bar

### Apple (□) menu

The first command of the Apple menu is the "About MacMul..." command. This command displays a dialog window (figure 2.2) containing some general information and in particular the number of the MacMul version being used. Click the OK button to continue

**MacMul: Multivariate data analysis  
Principal components analysis  
Correspondence analysis  
Multiple correspondence analysis**

**Version 3.01**

**Jean THIOULOUSE  
UNIVERSITE LYON 1  
LABORATOIRE DE BIOMETRIE - U.R.A. CNRS 243  
69622 VILLEURBANNE CEDEX - FRANCE**

**OK**

*Figure 2.2:  
Window called by the "About MacMul..." command.*

## **File menu**

The "File" menu offers some commands intended to make handling of the files used during analysis easier and some commands of general interest (figure 2.3).



*Figure 2.3:  
MacMul "File" menu.*



## Help

This command displays the on line help. The corresponding text is contained in the "MacMul.hlp" file which should be located in the folder where MacMul is found. If not, the program proposes looking for it in other folders or on another disk. The "Next" and "Stop" buttons allow passing to the next text page or returning to use of MacMul.

## Transfo. TEXT -> BIN

This command transforms a TEXT file into a binary file. After the text file to be transformed has been chosen (with the standard file selection window), the user must mark its number of columns, the column separators and the symbol of the decimal point (figure 2.4).

**Number of columns:**

**Separators:**

- ☒ **Spaces (ASCII 32):**
- ☒ **Comma (ASCII 44):**
- ☒ **Tabs (ASCII 9):**
- ☒ **Line feed (ASCII 10):**
- ☐ **Vert. tab. (ASCII 11):**
- ☐ **Form feed (ASCII 12):**
- ☐ **Carriage return (ASCII 13):**
- ☐ **Other: ASCII code=**

**Decimal mark:**

- ☒ **Period ( . )**
- ☐ **Comma ( , )**

**OK**

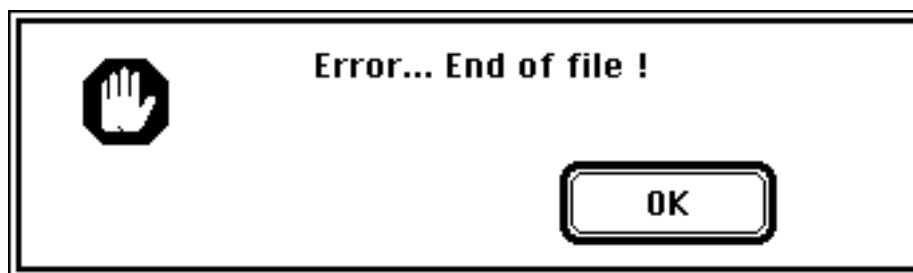
*Figure 2.4:  
Dialog box of the "Transfo. TEXT -> BIN" command.*

The text file must contain numbers, written with any format. The real numbers can be written with a comma as decimal symbol, e.g.: **3,14** (for French or gallicized program) or a full stop, e.g.: **3.14**.

The column separators can be chosen; the options presented by default (spaces, commas, tabulations and linefeeds) make it possible to read most of the text files found on the Macintosh. Two successive columns can be separated by any number of separators.

If the decimal point symbol is a comma, this character can obviously not be used as separator. A different separator from those proposed in the list can be chosen, by clicking the "Other" checkbox and typing its ASCII code in the corresponding field.

During the transformation MacMul displays the number of rows read. At the end, if the number of values contained in the file is not a multiple of the column number indicated, MacMul displays the following error message (figure 2.5):

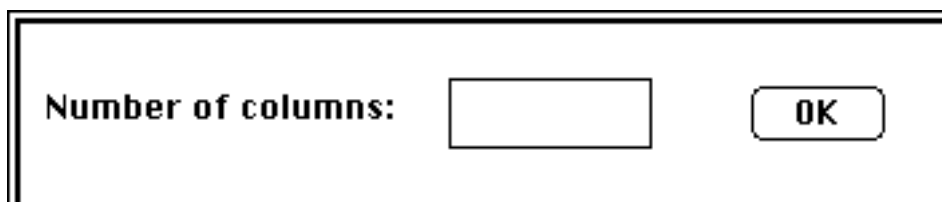


*Figure 2.5:  
End of file warning box.*

The contents of the resulting file is thus undefined...

### **Transfo. BIN -> TEXT**

This command corresponds to the inverse transformation of the preceding: it allows transformation of a binary file created by MacMul (or GraphMu, or ADECO), into a text file which could be used in a word processor or a spreadsheet. The text file obtained is in free format with tabs as separators and a decimal point (full stop, not a comma), as decimal symbol. MacMul asks the user to indicate the number of columns of the binary file (figure 2.6). One can thus give here a number of columns different from that declared when the file was created (but be careful with the resulting file structure in this case...).



*Figure 2.6:  
Window allowing indication of the number of columns  
of a binary file during transformation from binary to text.*

In the **GraphMu** program, the "Transfo BIN -> TEXT" command allows choice of a different output format, and specification of a FORTRAN writing format (the tabulation character is thus obtained by typing **option-t** ).

### **List BIN file**

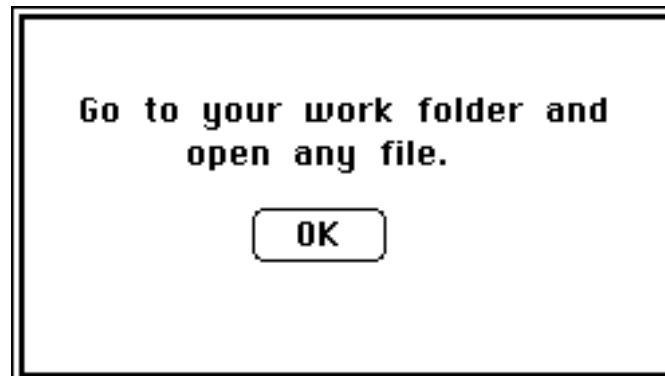
This command displays the contents of a binary file on the screen. Scientific notation is used, with 4 figures after the point (e.g.: 0.6544E-01, or -.5792E+03). Because of the window width, only 7 numbers per line can be displayed; if the lines contain more than 7 numbers, they are displayed on the following lines.

## **Print**

This command prints a text file. After the file has been selected the user has access to the two standard printing control dialog windows ("Page setup" and "Print" commands).

## **Work folder**

This command allows selection of a work folder (figure 2.7). It is only useful for continuing an analysis already begun in another folder, or for resuming an analysis after a program interruption.



*Figure 2.7:  
"Work folder" command dialog window.*

Using the standard file selection window one should select **any** file in the folder chosen and click the "Open" button.

## **Transpose file**

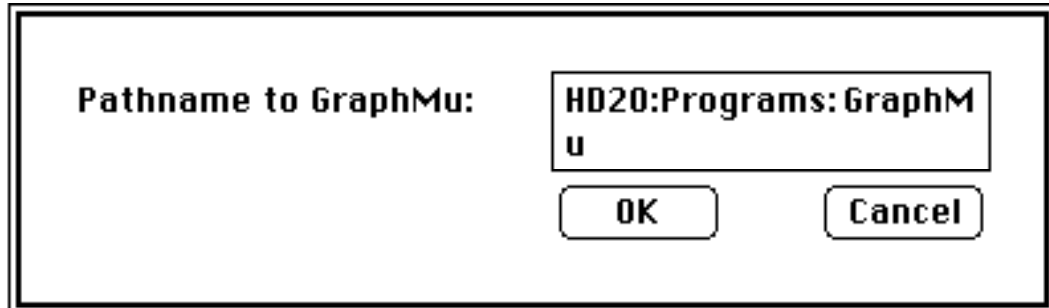
This command allows transposition of a file (the rows of the initial file become the columns of the new file). One must indicate the number of columns of the file to be transposed (figure 2.6). The name proposed for the transposed file is formed by adding the extension ".trn" to the name of the initial file. This operation is necessary before running the "Double discrimination" command of the CA menu.

## **Transfer**

Allows moving to another application without returning to the Finder.

## GraphMu

Allows going directly to the GraphMu program, without returning to the Finder and without changing the work folder. If MacMul cannot find GraphMu (e.g. if the GraphMu folder has been changed, or when MacMul is used for the first time on a Macintosh), a dialog box allows specification of the complete access path (figure 2.8).



*Figure 2.8:  
Dialog window allowing indication of the access path to GraphMu.*

This path will be written in MacMul resources and will then be available for use; it must be given under the form "HD20:Program:Mul:GraphMu" (first the disk name, followed by the folder names separated by a colon (:) and last the name of the GraphMu version used).

This command will **not** work when using Multifinder

## Quit

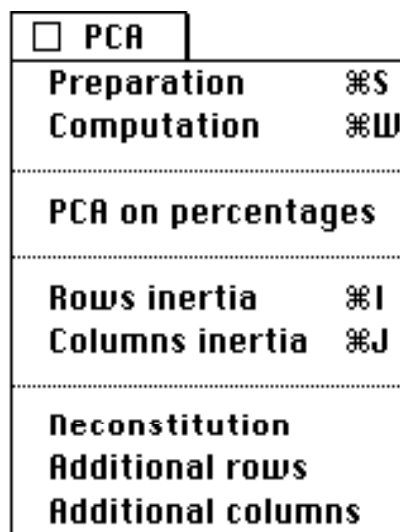
Returns to the Finder.

## Edit menu

The "Edit" menu is of no use with the MacMul program. It may be used with desk accessories (Apple menu).

### 3 - PCA MENU

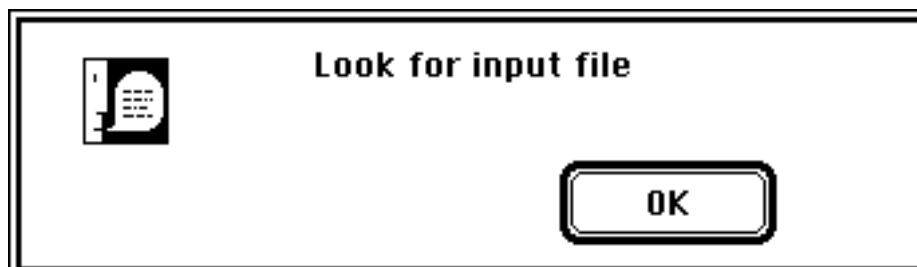
The PCA menu, like the CA and MCA menus, begins with the two commands "Preparation" and "Computation". One then finds interpretation aid commands: PCA on percentages (row profiles), inertia analyses of rows and columns, reconstitution of data and projection of additional rows and columns (figure 3.0).



*Figure 3.0:  
MacMul "PCA" menu*

#### Preparation

The program asks the user to choose the input file (figure 3.1). This file must be binary (use the "Transfo. TEXT -> BIN" command of the "File" menu).



*Figure 3.1:  
First dialog box of the "Preparation" command*

A dialog window (figure 3.2) then allows specification of the number of rows and columns of this file, as well as the analysis title. This title is a generic name which will be used to create the output file names automatically.

The type of PCA (centred PCA, normalized PCA, non-centred PCA, or general PCA of a triplet with two diagonal norms), must be chosen using a series of "radio buttons".

Number of rows ? 24

Number of columns ? 10

Title of analysis ? cn1

☒ Centered PCA

☐ Standardized PCA

☐ Non centered PCA

☐ PCA (X, Dp, Dn)

OK Cancel

Figure 3.2:  
Main dialog window of a PCA.

Two files are created during this first stage.

The first has a name corresponding to the title of the analysis without extension (here, e. g., "**cn1**"); depending on the type of PCA chosen, it contains centered data (centered PCA), normalized data (normalized PCA), or simply the initial data (non-centered PCA).

The second file is called "**cn1.CPMA**" and contains a description of the analysis: number of rows, number of columns, code of the type of PCA (centered = 1, normalized = 2, non-centered = 3 and general = 4), total inertia, means and variances of each variable.

If a general PCA is to be run (button PCA [X, Dp, Dn] of the dialog window), before using the "Computation" command, one must have created two files in the work folder which must be called "**xxx.CPPL**" and "**xxx.CPPC**" (xxx represents the analysis title). These files must contain the row and column weightings. These two files must be binary and must have only one column.

The "**xxx.CPPL**" file must have a number of **rows** equal to the number of **rows** of the data file.

The "**xxx.CPPC**" file must have a number of **rows** equal to the number of **columns** of the data file.

## Computations

This second phase runs the main calculations of a classical PCA (calculation of the covariance or correlation matrix, diagonalisation, calculation of the factorial coordinates of rows and columns).

The user must first indicate the title of the analysis for which he wishes to perform the calculations (figure 3.3). It is possible to store the matrix which is to be diagonalized in a file. According to the PCA type, this matrix will be the covariance matrix (centered PCA), the correlation matrix (standardized PCA) or some scalar products matrix (non-centered or general PCA).



**Title of analysis ?**

☒ **Store scalar products matrix**

**OK** **Cancel**

*Figure 3.3:  
First dialog window of the "Computation" part of a PCA.*

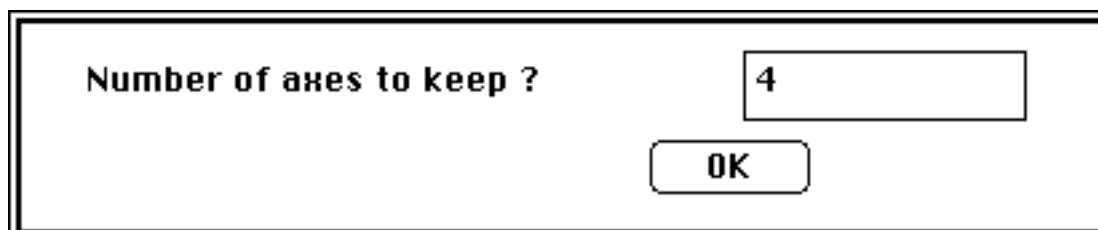
After a longer or shorter time, depending on the size of the data file, the list of eigenvalues with the corresponding inertia percentages is displayed (figure 3.4). The first line indicates what the different columns correspond to: first the order number of the eigenvalue ("Num."), then its value ("Eigenval."), the corresponding inertial percentage ("%Iner.") and finally the cumulative percentage ("R.Sum"). The display is by groups of two values on each line, and only the first 24 values are displayed.

Num.	Eigenval.	%Iner.	R.Sum	Num.	Eigenval.	%Iner.	R.Sum
EV 1	:0.5745E+01	.5745	.5745	EV 2	:0.1430E+01	.1430	.7175
EV 3	:0.1084E+01	.1084	.8259	EV 4	:0.6761E+00	.0676	.8935
EV 5	:0.5244E+00	.0524	.9459	EV 6	:0.3033E+00	.0303	.9763
EV 7	:0.1462E+00	.0146	.9909	EV 8	:0.5371E-01	.0054	.9963
EV 9	:0.2273E-01	.0023	.9985	EV 10	:0.1454E-01	.0015	1.0000

*Figure 3.4:  
Display of the list of eigenvalues.*

As an example, in figure 3.4, one sees that the first eigenvalue (EV 1) is 5.745; it corresponds to an inertia percentage of 57.45%. The second eigenvalue (EV 2) is 1.430; it corresponds to an inertia percentage of 14.3%. The sum of the first two percentages is 71.75%.

At this point a window allows choice of the number of factors for which the program must calculate the factorial coordinates (figure 3.5).



*Figure 3.5:  
Dialog window for the number of factors to keep in coordinate files.*

During this stage five files are created:

the "**cn1.CPLI**" file contains the factorial coordinates of the rows,

the "**cn1.CPCO**" file contains the factorial coordinates of the columns,

the "**cn1.CPCL**" file contains the row contributions to the trace (useful for the inertia analysis),

the "**cn1.CPCC**" file contains the column contributions to the trace (useful for the inertia analysis),

the "**cn1.CPVP**" file contains the number of saved factors, the total inertia and the list of all eigenvalues arranged in decreasing order.

Moreover, the "**cn1.CPDI**" file may contain the matrix which was to be diagonalized (e.g. the correlation matrix if one has performed a standardized PCA).

The following commands correspond to interpretation aids and are optional. The "**cn1.CPLI**" and "**cn1.CPCO**" files can immediately be used with GraphMu to draw the graphic representations (factorial planes).

## **Row inertia**

This command allows the inertia analysis of rows (to be used only after the "Preparation" and "Computation" commands). A dialog window opens to ask for the title of the analysis for which the inertia calculations are to be performed. A second window then displays the number of factors which were saved during the "Computation" phase and asks for the number of factors to be processed (figure 3.6).



Number of axes kept: 4

Number of axes to analyse ? 3

OK

Figure 3.6:  
Dialog window for the number of factors to be processed

The number of factors to be processed must be less than or equal to the number of factors saved in the analysis. One must then choose the output file, which will contain the text describing the inertia analysis (figure 3.7).

Choose output file.

OK

Figure 3.7:  
Dialog window for the output file.

The standard file selection window appears and proposes "cn1.CPIL" (for **Line Inertias**) by default as name for this file. It will contain a description of the inertia analysis, showing as lines the data file rows and as columns the absolute and relative contributions of the rows to the factors to be processed, as well as their residual contributions, the weights and the contributions to the trace.

The **absolute** contributions indicate that part of the **variability of a factor which can be imputed to an element** (row or column).

The **relative** contributions indicate that part of the **variability of an element** (row or column) **accounted for by a factor**.

The example below shows the beginning of a row inertia analysis (analysis of the first five rows only). In the heading of the description, one can see that the title of the analysis is "cn1", that the data file has 24 rows and 10 columns and that four factors were saved. Only the first three were the subject of the inertia analysis:

PCA: Inertia analysis of rows.  
Title of analysis:  
cn1 24 rows 10 columns 4 factors

Absolute contributions of rows.  
Factors are in columns.

1	179.	2.	400.
2	42.	40.	683.
3	144.	461.	11.
4	40.	25.	89.
5	116.	376.	424.

etc.

The first column gives the number of rows and the next three are the **absolute** contributions of each row to the first three factors. In this table, the sum of the terms of each column is equal to 10 000.

The remainder of the file contains the **relative** contributions:

Relative contributions of rows.  
Factors are in columns, followed by  
residual contributions, weights, and  
contributions to the trace

1	4566.	11.	1925.		3498.		417.		225.
2	1433.	345.	4439.		3783.		417.		167.
3	3146.	2509.	46.		4298.		417.		263.
4	1029.	158.	434.		8379.		417.		223.
5	1870.	1512.	1294.		5324.		417.		355.

etc.

The first column gives the number of the rows and the three following columns are the relative contributions of each row to the first 3 factors. Column 5 corresponds to the residual contributions, i.e. the part of variability taken into account by the other factors (factors 4 to 10). The last two columns (6 and 7) indicate the weight of the rows and their contribution to the trace respectively.

In this table, the sum of relative contributions and the residual contribution for each row is 10 000.

This file is a text file which can be read by any word processing program (to keep the column alignment, one must use a font with non-proportional spacing, such as Monaco or Courier).

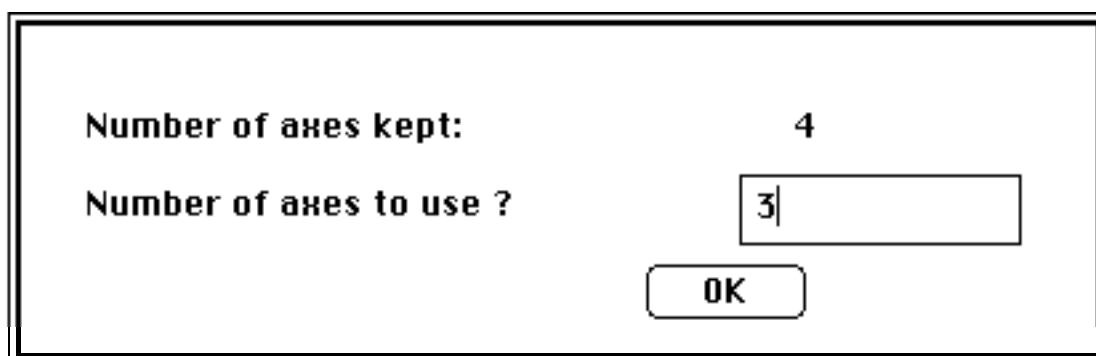
## Column inertia

This command is identical to that for row inertia analysis. The dialog is the same except for the name proposed by default for the output file, which here is "cn1.CPIC" (for **C**olumn **I**nter**I**as). The contents of this file is constructed following the same model as the previous one.

## Reconstitution

This command allows a reconstitution of the data, starting from one or several factors. The reconstituted data are always the data processed by the analysis, i.e. the centered data in the case of a centered PCA, the normalized data in the case of a normalized PCA and the raw data in the case of a non-centered PCA.

A first window asks for the analysis title for which the data is to be reconstituted. A second window then displays the number of factors which were saved in this analysis and asks for the number of factors to be used for the reconstitution (figure 3.8).



The dialog window has a double-line border. It contains two labels: "Number of axes kept:" and "Number of axes to use ?". To the right of the first label is the number "4". To the right of the second label is a text input field containing the number "3". Below the input field is an "OK" button.

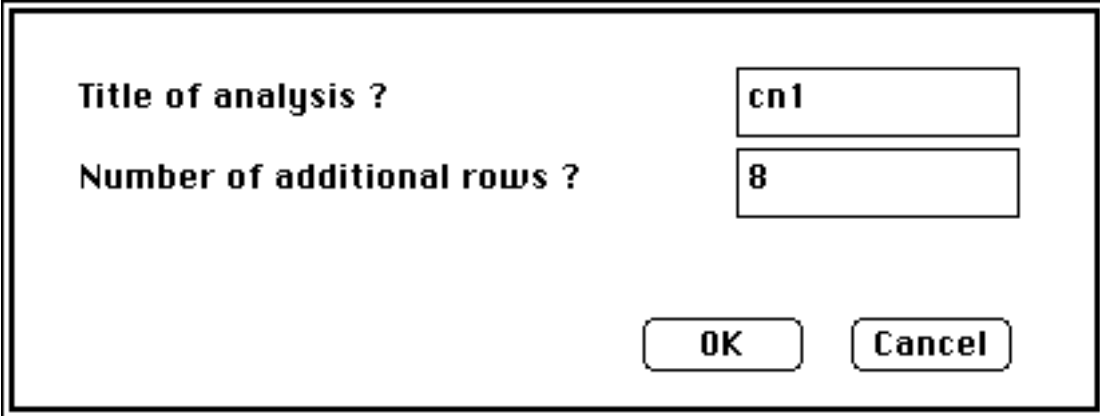
*Figure 3.8:  
Dialog window for data reconstitution*

The number of files created will be equal to this number of factors used for the reconstitution. If one uses three factors, these files will be called "**cn1.CPR1**", "**CN1.CPR2**" and "**cn1.CPR3**".

The "**cn1.CPR1**" file will contain the data reconstituted with the first factor, the "**CN1.CPR2**" file will contain the data reconstituted with the first **and** second factors and the "**cn1.CPR3**" file will contain the data reconstituted with the first, second **and** third factors. All these files have the same number of rows and columns as the initial data file.

## Additional rows

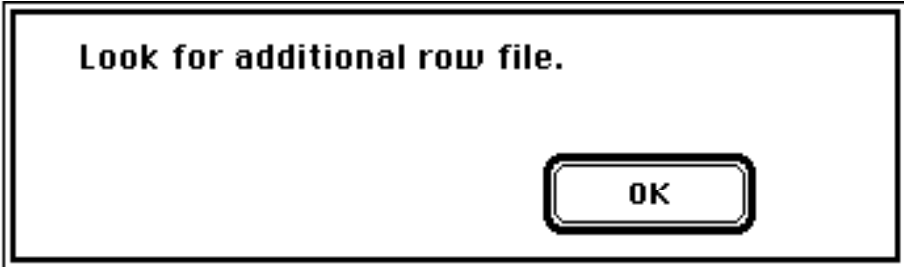
This command allows projection of additional rows in a PCA. A first window asks for the analysis title into which the additional rows are to be projected and the number of rows to project (figure 3.9).



A dialog window with a double border. It contains two labels on the left: "Title of analysis ?" and "Number of additional rows ?". To the right of each label is a text input field. The first field contains the text "cn1" and the second field contains the text "8". At the bottom right of the window are two buttons: "OK" and "Cancel".

*Figure 3.9:  
Dialog window for the projection of additional rows.*

Using the standard file selection window one must then indicate the file containing the rows to be projected. This file must be a binary file and it must have **the same number of columns** as the initial analysis file into which the projection is to be made (figure 3.10).



A dialog window with a double border. It contains a single label: "Look for additional row file.". At the bottom right of the window is a button labeled "OK".

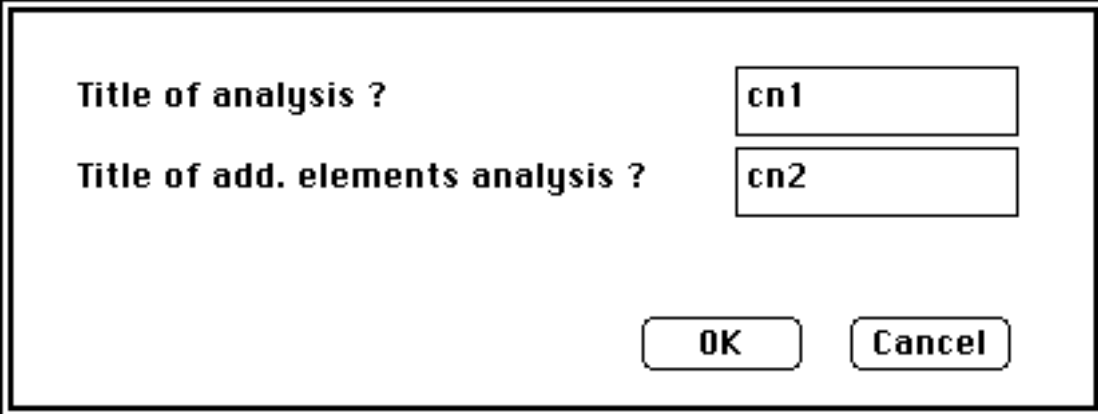
*Figure 3.10:  
Dialog window for the file of rows to be projected.*

The coordinates of additional rows are automatically calculated for all the factors which were saved in the analysis ("Computation" stage).

The output file takes the name of the file containing the additional rows, followed by the extension ".CPLS"; it contains the factorial coordinates of the additional rows and thus has a number of **rows** equal to the number of **rows** of the additional row file. Its number of **columns** is equal to the number of **factors** saved in the analysis where the rows were projected.

## Additional columns

This command allows projection of additional columns in a PCA. The additional column file **must have passed through the "Preparation"** step of the PCA menu. The first window asks for the analysis title into which the projection is to be made and the additional column analysis title (figure 3.11).



The dialog window has a double-line border. It contains two labels on the left: "Title of analysis ?" and "Title of add. elements analysis ?". To the right of each label is a text input field. The first input field contains the text "cn1" and the second contains "cn2". At the bottom right of the dialog are two buttons: "OK" and "Cancel".

*Figure 3.11:  
Dialog window for the projection of additional columns*

These two analyses must be of the **same type** (non-centered, centered or normalized) and must also have the **same number of rows**. If not, the program displays an error message and does not perform the calculations.

If the two analyses are compatible, the output file will have the name of the analysis title of the additional columns followed by the extension **".CPCS"** and will contain the factorial coordinates of the additional columns. It will thus have a number of **rows** equal to the number of **columns** of the additional column file and a number of **columns** equal to the number of **factors** saved in the analysis **where the projection was made**.

## PCA on profiles

In the special case where the data file rows represent distributions (the sum of each row is therefore equal to 1) this command, used after the "Preparation" (choosing obligatorily a centered PCA) and "Computation" commands, allows a simultaneous representation of rows and columns. In the special case where there are only three columns one thus obtains, rotated, the classical triangular representation of granulometry.

The program just asks for the analysis title for which the calculations must be performed (this title was fixed during execution of the "Preparation" command).

Two files are created:

the "**cn1.CPPO**" file contains as many rows as are in the data file plus one (the first row gives the position of the center of gravity), and as many columns as there were saved factors in the "Computation" command,

the "**cn1.CPVE**" file contains the coordinate of columns; it has the same size (rows and columns) as the "cn1.CPCO" file. The graphical representations made with the "cn1.CPVE" and "cn1.CPPO" files can be superimposed (at the same scale).

## 4 - CA MENU

The CA menu also begins with the two commands "Preparation" and "Computation". One then finds the same interpretation aid commands as in PCA: inertia analyses of rows and columns, reconstitution of data and projection of additional rows and columns. "Double discrimination" is an extra command (figure 4.0).

<input type="checkbox"/> CA	
<b>Preparation</b> ⌘E	
<b>Computation</b> ⌘D	
<hr/>	
<b>Rows inertia</b> ⌘L	
<b>Columns inertia</b> ⌘K	
<hr/>	
<b>Reconstitution</b>	
<b>Additional rows</b>	
<b>Additional columns</b>	
<hr/>	
<b>Double discrimination</b>	

*Figure 4.0:  
MacMul "CA" menu*

### Preparation

This first stage is similar to the corresponding stage in PCA. The main dialog window is slightly different however (figure 4.1).

<b>Number of rows ?</b>	<input type="text"/>
<b>Number of columns ?</b>	<input type="text"/>
<b>Title of analysis ?</b>	<input type="text"/>
<div><input type="button" value="OK"/> <input type="button" value="Cancel"/></div>	

*Figure 4.1:  
Main dialog window of a CA*

Four files are created during this first stage:

a file whose name corresponds to the analysis title without extension (e.g. "**xxx**"), containing the data processed,

a file called "**xxx.FCMA**" containing a description of the analysis (number of rows, columns and individuals),

two files called "**xxx.FCPL**" and "**xxx.FCPC**" which contain the weightings of rows and columns.

## Computations

This second phase runs the main CA calculations (calculation of the matrix to diagonalise, diagonalisation, calculation of the factorial coordinates of the rows and columns). The list of eigenvalues with the corresponding inertia percentages is displayed (cf. figure 3.4), then a dialog window allows choice of the number of factors to save (figure 3.5). The factorial coordinates will be calculated on these factors only.

Five or six files are created:

the "**xxx.FCLI**" file contains the factorial coordinates of the rows (**lines**),

the "**xxx.FCCO**" file contains the factorial coordinates of the **columns**,

the "**xxx.FCVP**" file contains the number of factors saved, the total inertia and the list of eigenvalues arranged in decreasing order.

the "**xxx.FCCL**" and "**xxx.FCCC**" files contain the contributions to the trace of rows and columns.

the "**xxx.FCDI**" file may contain the matrix which was to be diagonalized.

Only the "**xxx.FCLI**" and "**xxx.FCCO**" files are needed to draw the factorial diagrams with GraphMu.

## Row inertia

This command allows the inertia analysis of the rows of the file (to be used only after the "Preparation" and "Computation" commands). A dialog window opens to ask for the title of the analysis for which the inertia calculations are to be performed.

As in PCA, a second window displays the number of factors which were saved and asks for the number of factors to be processed (figure 3.6). The number of factors to be processed must be less than or equal to the number of factors saved in the analysis. One must then choose the output file, which will contain the text describing the inertia analysis. The standard file selection window appears and proposes "**xxx.FCIL**" by default as name for this file.



This file will contain a description of the inertia analysis, showing in the line the data file rows and in the columns their absolute and relative contributions to the factors to be processed, as well as the residual contributions, the weights and the contributions to the trace.

The presentation and meaning of the different elements are the same as in PCA.

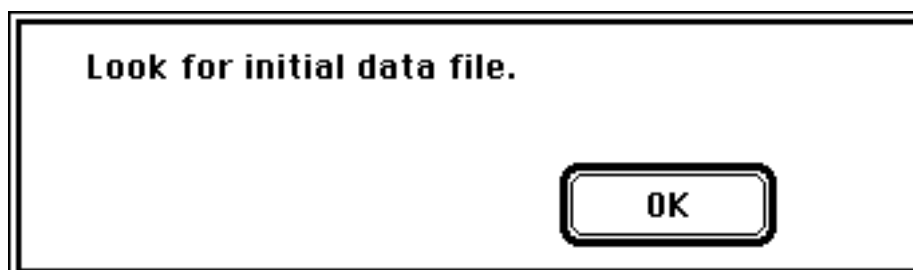
## Column inertia

This command is identical to that for row inertia analysis. The dialog is the same except for the name proposed by default for the output file which here is "**xxx.FCIC**".

## Reconstitution

This command allows a reconstitution of the data from one or several factors. The reconstituted data correspond to the raw data (initial data file).

A first window asks for the analysis title for which the data is to be reconstituted. A second window then displays the number of factors which were saved in this analysis and asks for the number of factors to be used for the reconstitution (figure 3.8). The program then asks to look for the initial analysis data file (figure 4.2) with the standard file selection window.



*Figure 4.2:  
Dialog window for the initial file.*

Unlike the "Reconstitution" command of the PCA menu, there is only one output file, called "**xxx.FCRE**". When one asks for a reconstitution with several factors, this file contains the reconstitution with all the factors required. It is the same size (rows and columns) as the initial data file.

## Additional rows

This command allows projection of additional rows in a CA. A first window asks for the analysis title into which the additional rows are to be projected and the number of rows to project.

Using the standard file selection window one must then indicate the file containing the rows to be projected. This file must be a binary file and it must have the same number of **columns** as the analysis into which the projection is to be made.

The coordinates of additional rows are calculated for all the factors which were saved in the analysis. The output file takes the name of the file containing the additional rows, followed by the extension ".FCLS"; it contains the factorial coordinates of the additional rows and thus a number of **columns** equal to the number of **factors** saved in the analysis where the rows were projected. Its number of **rows** is equal to the number of **additional rows**.

## Additional columns

This command allows projection of additional columns in a CA. The first window asks for the analysis title into which the projection is to be made and the number of columns to project.

One must then, using the standard file selection window, indicate the file containing the columns to be projected. This file must be binary, and it must have the same number of **rows** as the analysis into which the projection is to be made.

The additional column coordinates will be calculated on all the factors which were saved in the analysis. The output file will have the name of the file containing the additional columns followed by the extension ".FCCS", it contains the factorial coordinates of the additional columns and will have a number of **columns** equal to the number of **factors** saved in the analysis where the projection was made. Its number of **rows** is equal to the number of **additional columns**.

## Double discrimination

This command allows interpretation of the results of a CA like those of a double discrimination analysis on the rows and columns of the contingency table, or like those of a generalised canonic analysis on the complete disjunctive tables associated with the two qualitative variables of the contingency table. One can thus obtain a coding for each of the cells of the contingency table.

A first window asks for the analysis title for which the calculations are to be performed. A second window displays the number of factors saved in this analysis and asks for the number of factors to be used for the double discrimination (figure 3.8). One must then look for the initial data file (figure 4.2) and the corresponding transposed file (obtained e.g. with the "Transpose file" command of the "File" menu) (figure 4.3).



Figure 4.3:  
Dialog window for the transposed file

Three files are automatically created: "**xxx.FCCA**", "**xxx.FCVL**" and "**xxx.FCVC**".

The "**xxx.FCCA**" file contains the factorial coordinates of the cells of the contingency table. The coordinate  $X_k(i,j)$  of cell  $(i,j)$  on axis  $k$  is given by:

$$X_k(i,j) = \left( \frac{1}{\sqrt{2\lambda_k\mu_k}} \right) (L_k(i) + C_k(j))$$

$L_k(i)$  and  $C_k(j)$  are the factorial coordinates of the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column on axis  $k$ ,  $\lambda_k$  is the  $k^{\text{th}}$  eigenvalue of the CA and  $\mu_k = 1 + R(\lambda_k)$

The "**xxx.FCCA**" file has a number of **rows** equal to the **product** of the number of rows by the number of columns of the initial file. It has a number of **columns** equal to the number of **factors** used, **plus 1**. This extra column (which is the last column of the file) contains the values of the cells of the scrolled contingency table. This allows a graphical representation of this file with the "Squares map" (or "Circles map") commands of GraphMu.

The "**xxx.FCVL**" and "**xxx.FCVC**" files contain the means-variances-covariances of the factorial coordinates of rows (or columns), weighted by the entries of the contingency table. These two files are of the "MVC" type used by GraphMu to draw ellipses or Gauss curves. If  $F$  is the number of factors used, they have  $2.F + F.(F-1)/2$  columns. The first  $F$  columns contain the means and the next contain the lower triangle of the covariance matrix read line by line. The factorial maps obtained with the "**xxx.FCCA**" file and the ellipses of inertia obtained with "**xxx.FCVL**" and/or "**xxx.FCVC**" can be superimposed.



## 5 - MCA MENU

The MCA menu has the two basic commands "Preparation" and "Computation". The aid to interpretation commands are: inertia analysis of rows and columns and projection of additional rows and columns. The data reconstitution command is replaced by the calculation and edition of variable-factor correlation ratios (figure 5.0).

<input type="checkbox"/> MCA	
Preparation	⌘T
Computation	⌘F
<hr/>	
Rows inertia	⌘M
Columns inertia	⌘N
Correlation ratios	⌘R
<hr/>	
Additional rows	
Additional columns	

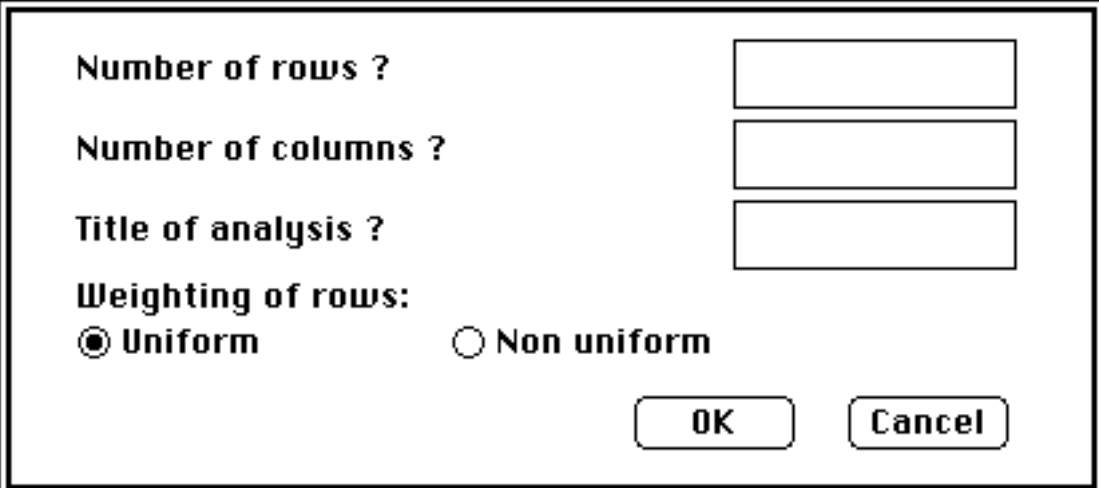
Figure 5.0:  
MacMul "MCA" menu

### Preparation

The program asks the user to choose the input file (figure 3.1). This file must contain **qualitative** variables, the categories being numbered from **1 to N**, without missing category (i.e. no category with no individual belonging to it).

A dialog window then allows the user to give the number of rows and columns of this file and the analysis title and to specify whether the row weighting is uniform (equal to 1) or not (figure 5.1).

A non-uniform weighting can allow grouping of individuals belonging to the same categories (for all the variables) in just one data file row. In this case, one must then give the name of the file which contains, for each data file row, the number of corresponding individuals (figure 5.2). This file must be a binary file, have one column and a number of rows identical to that of the data file.



Number of rows ?

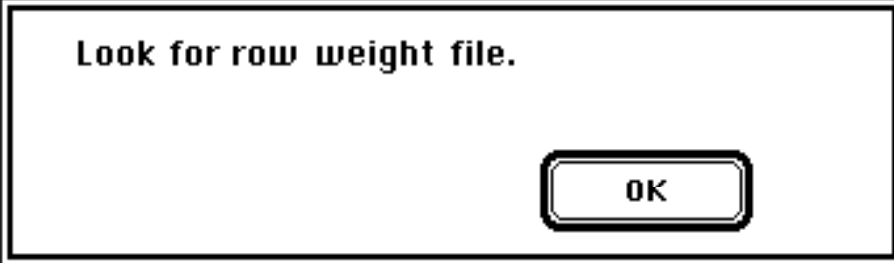
Number of columns ?

Title of analysis ?

Weighting of rows:

☒ Uniform ☐ Non uniform

*Figure 5.1:  
Main dialog window of a MCA*



**Look for row weight file.**

*Figure 5.2:  
Dialog window of the row weighting file search*

Four files are created during this first stage:

a file whose name corresponds to the analysis title without extension (e.g. "xxx"), containing the data processed by the analysis (weighted complete disjunctive table),

a file called "xxx.MCMA" containing a description of the analysis (number of rows, categories, variables and individuals, categories by variable and individuals by category),

two files called "xxx.MCPL" and "xxx.MCPC" which contain the weights of rows and categories.

## Computations

This second phase runs the main calculations of a MCA (calculation of the matrix to be diagonalised, diagonalisation, calculation of the factorial coordinates of the rows and columns (categories)). The list of eigenvalues with the corresponding inertia

percentages is displayed (cf. figure 3.4), then a dialog window allows choice of the number of factors to save (figure 3.5).

Five or six files are created:

the "**xxx.MCLI**" file contains the factorial coordinates of rows,

the "**xxx.MCCO**" file contains the factorial coordinates of categories,

the "**xxx.MCVP**" file contains the number of factors saved, the total inertia and the list of eigenvalues arranged in decreasing order.

the "**xxx.MCCL**" and "**xxx.MCCC**" file contain the contributions to the trace of rows and categories respectively.

the "**xxx.MCDI**" file may contain the matrix which was to be diagonalized.

Only the "**xxx.MCLI**" and "**xxx.MCCO**" file are needed to draw the factorial diagrams with GraphMu.

## Row inertia

This command allows the inertia analysis of rows (to be used only after the "Preparation" and "Computations" commands). A dialog window opens to ask for the title of the analysis for which the inertia calculations are to be performed.

As in PCA and CA, a second window displays the number of factors which were saved and asks for the number of factors to be processed. The number of factors to be processed must be less than or equal to the number of factors saved in the analysis. One must then choose the output file, which will contain the text describing the inertia analysis. The standard file selection window appears and proposes "**xxx.MCIL**" by default as name for this file.

This file will contain a description of the inertia analysis, showing in line the data file rows and in column their absolute and relative contributions to the factors to be processed, as well as the residual contributions, the weights and the contributions to the trace.

The presentation and meaning of the different analysis elements are the same as in PCA and in CA.

## Columns inertia

This command is identical to that for row inertia analysis. The dialog is the same except for the name proposed by default for the output file which here is "**xxx.MCIC**".

## Correlation ratios

This command allows study of the variable-factor correlation ratios. A first window asks for the analysis title for which the study is to be performed. A second window then displays the number of factors saved in this analysis and asks for the number of factors to be processed. One must then choose the output file, which will contain the description of the analysis. The standard file selection window is displayed and proposes as name by default for this file "**xxx.MCRC**". It will contain a description of the variables (number of categories and number of individuals per category) and a table of the variable-factor correlation ratios (variable as line and factors as columns).

The example below shows the beginning of one of these files. In the heading one can see that the input file had 24 rows and 3 variables, the 3 variables having 10 categories in all.

In the categories description, one can see, e.g., that the first variable has 3 categories and that 15 individuals (i.e. 62.5%) belong to the first category, 7 (i.e. 29.17%) to category 2 and 2 (i.e. 8.33%) to category 3.

In the correlation ratios table one finds as lines the 3 variables and as columns the 4 factors. The correlation ratio of variable 3 with factor 2 is, e.g., 0.7824.

MCA: Correlation ratios.

Title of analysis:

cm1 24 rows 3 variables 10 categories

Description of categories:

Variable number 1 has 3 categories:

1 | Cat.: 1 Ind.: 15 Freq.: .6250  
2 | Cat.: 2 Ind.: 7 Freq.: .2917  
3 | Cat.: 3 Ind.: 2 Freq.: .0833

Variable number 2 has 2 categories:

4 | Cat.: 1 Ind.: 8 Freq.: .3333  
5 | Cat.: 2 Ind.: 15 Freq.: .6250  
6 | Cat.: 3 Ind.: 1 Freq.: .0417

Variable number 3 has 3 categories:

7 | Cat.: 1 Ind.: 4 Freq.: .1667  
8 | Cat.: 2 Ind.: 3 Freq.: .1250  
9 | Cat.: 3 Ind.: 11 Freq.: .4583  
10 | Cat.: 4 Ind.: 6 Freq.: .2500

Variables/Factors correlation ratios.

Factors are in columns.

VAR 1 | .4199 | .2642 | .4081 | .4692 |  
VAR 2 | .7582 | .4598 | .0898 | .3932 |  
VAR 3 | .7053 | .7824 | .7220 | .1297 |



## Additional rows

This command allows projection of additional rows in a MCA. The additional rows file **must have passed through the "Preparation"** command of the MCA menu.

A first window asks for the analysis title into which the additional rows are to be projected and the analysis title of the additional rows (figure 3.11). These two analyses must have the **same number of categories**.

The name of the output file will be the analysis title of the additional rows followed by the extension **".MCLS"** and will contain the factorial coordinates of the additional rows. It will therefore have the same number of **rows** as the additional rows file and a number of **columns** equal to the number of **factors** saved in the analysis into which the projection has been made.

## Additional columns

This command allows projection of additional **qualitative variables** in a MCA. The additional qualitative variable file **must have passed through the "Preparation"** command of the MCA.

A first window asks for the analysis title into which the projection is to be made and the analysis title of the additional qualitative variables (figure 3.11). These two analyses must have the **same number of rows**.

The name of the output file will be the analysis title of the additional qualitative variables followed by the extension **".MCCS"** and will contain the factorial coordinates of the additional categories. It will therefore have a number of **rows** equal to the **total number of categories** of the extra variables and a number of **columns** equal to the number of **factors** saved in the analysis into which the projection has been made.



## 6 - EXAMPLES OF USE

### PCA

We will use a data set taken from a paper by Doledec and Chessel (Doledec, S. and Chessel, D., 1987. Rythmes saisonniers et composantes stationnelles en milieu aquatique. I - Description d'un plan d'observation complet par projection de variables. *Acta OEcological, OEcol. Gener.*, **8**, 403 - 426), corresponding to the table below (table 2). The 10 columns correspond to 10 physico-chemical variables recorded at 6 stations situated along a stream, at 4 dates.

10.00	41.00	8.50	295.00	110.00	2.30	1.40	.12	3.40	.11
13.00	62.00	8.30	325.00	95.00	2.30	1.80	.11	3.00	.13
1.00	25.00	8.40	315.00	91.00	1.60	.50	.07	6.40	.03
3.00	118.00	8.00	325.00	100.00	1.60	1.20	.17	1.80	.19
11.00	158.00	8.30	315.00	13.00	7.60	3.30	2.85	2.70	1.50
13.00	80.00	7.60	380.00	20.00	21.00	5.70	9.80	.80	3.65
3.00	63.00	8.00	425.00	38.00	36.00	8.00	12.50	2.20	6.50
3.00	252.00	8.30	360.00	100.00	9.50	2.90	2.52	4.60	1.60
11.00	198.00	8.50	290.00	113.00	3.30	1.50	.40	4.00	.10
15.00	100.00	7.80	385.00	46.00	15.00	2.50	7.90	7.70	4.50
2.00	79.00	8.10	350.00	84.00	7.10	1.90	2.70	13.20	3.70
3.00	315.00	8.30	370.00	100.00	8.70	2.80	2.80	4.80	2.85
12.00	280.00	8.60	290.00	126.00	3.50	1.50	.45	4.00	.73
16.00	140.00	8.00	360.00	76.00	12.00	2.60	4.90	8.40	3.45
3.00	85.00	8.30	330.00	106.00	2.00	1.40	.42	12.00	1.60
3.00	498.00	8.30	330.00	100.00	4.80	1.60	1.04	4.40	.82
13.00	322.00	8.50	285.00	117.00	3.60	1.60	.48	4.60	.84
15.00	160.00	8.40	345.00	91.00	1.70	1.90	.22	10.00	1.74
2.00	72.00	8.60	305.00	91.00	1.60	.90	.10	9.50	1.25
2.00	390.00	8.20	330.00	100.00	1.70	1.20	.56	5.00	.60
11.00	303.00	8.50	245.00	100.00	1.70	.90	.05	2.70	.16
13.00	310.00	8.20	285.00	82.00	8.50	1.60	.59	3.70	.60
4.00	181.00	8.60	270.00	105.00	2.80	.50	.10	3.66	.43
3.00	480.00	8.20	290.00	100.00	1.30	.80	.04	2.20	.13

Table 2:  
Contents of the "Meau" file.

The first operation thus consists of typing this data set into a Macintosh program (Edit, Excel, ...) and of recording it in a text file. This file is called "Meau.txt". It must then be translated into binary with the "Transfo. TEXT --> BIN" command of the "File" menu of MacMul or GraphMu. The resulting file is the "Meau" file. The information for these files is given in figure 6.1.

To run the PCA, one now needs just use MacMul and select the "Preparation" command of the "PCA" menu. After the Meau file has been selected, its number of rows and columns (figure 3.2) must be indicated in the main dialog window. If a

normalized PCA is to be run, as in this example, the corresponding button must be selected. The title of the analysis chosen here is "cn1". The files created during this first stage are shown in figure 6.2..

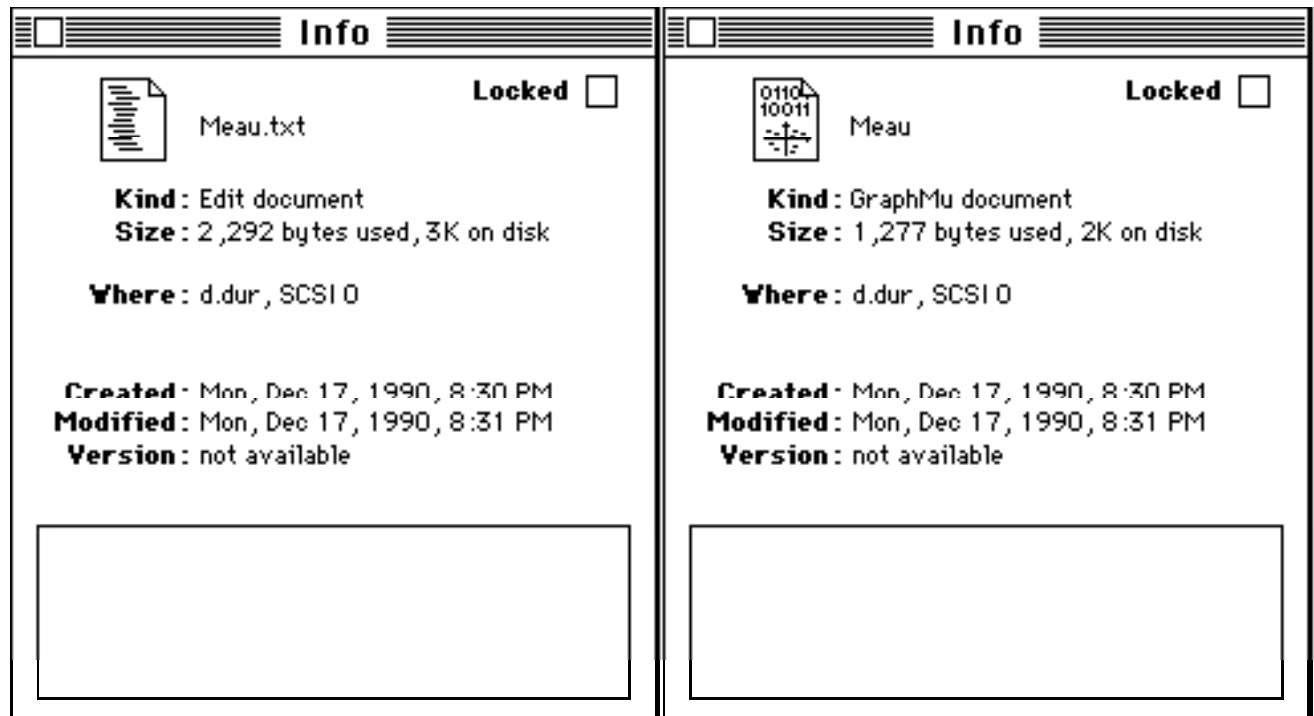


Figure 6.1:  
Information windows of the "Meau.txt" and "Meau" files.  
The size of the "Meau.txt" file can vary depending on how it is typed in.

The "cn1" file can be translated into a text file with the "Transfo. BIN --> TEXT" command of the "File" menu of MacMul or of GraphMu. The contents of this file is shown below (there are only 3 figures after the comma, using the possibility of using a **FORTRAN** format with the GraphMu transformation command). Values are normalized by variable (the normalisation consists of subtracting the mean and dividing by the standard deviation).

One can check that the mean per column of this table is zero and that the variance is 1 (to within numerical errors of approximation).

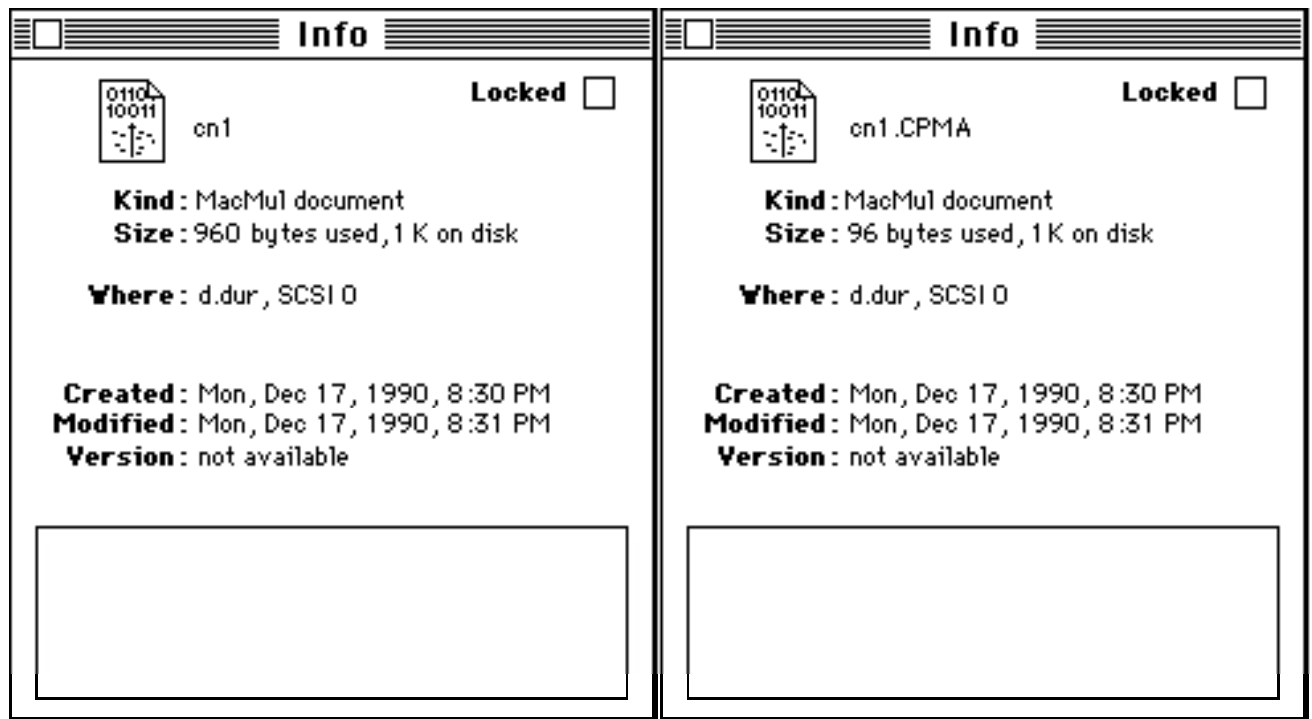


Figure 6.2:  
Information windows of the "cn1" and "cn1.CPMA" files

+0.439	-1.148	+0.914	-0.735	+0.778	-0.564	-0.416	-0.604	-0.559	-0.864
+1.013	-0.993	+0.116	+0.000	+0.255	-0.564	-0.173	-0.607	-0.684	-0.852
-1.284	-1.266	+0.515	-0.245	+0.116	-0.654	-0.965	-0.619	+0.374	-0.912
-0.901	-0.579	-1.080	+0.000	+0.429	-0.654	-0.538	-0.589	-1.057	-0.816
+0.630	-0.283	+0.116	-0.245	-2.600	+0.113	+0.741	+0.220	-0.777	-0.030
+1.013	-0.860	-2.675	+1.347	-2.356	+1.826	+2.203	+2.320	-1.368	+1.259
-0.901	-0.985	-1.080	+2.449	-1.729	+3.743	+3.604	+3.135	-0.932	+2.968
-0.901	+0.411	+0.116	+0.857	+0.429	+0.356	+0.497	+0.121	-0.186	+0.030
+0.630	+0.012	+0.914	-0.857	+0.882	-0.437	-0.355	-0.520	-0.373	-0.870
+1.396	-0.712	-1.877	+1.470	-1.451	+1.059	+0.254	+1.746	+0.778	+1.769
-1.093	-0.867	-0.681	+0.612	-0.128	+0.049	-0.112	+0.175	+2.488	+1.289
-0.901	+0.877	+0.116	+1.102	+0.429	+0.253	+0.437	+0.205	-0.124	+0.779
+0.822	+0.618	+1.312	-0.857	+1.335	-0.411	-0.355	-0.505	-0.373	-0.492
+1.587	-0.416	-1.080	+0.857	-0.406	+0.675	+0.315	+0.840	+0.995	+1.139
-0.901	-0.823	+0.116	+0.122	+0.638	-0.603	-0.416	-0.514	+2.115	+0.030
-0.901	+2.230	+0.116	+0.122	+0.429	-0.245	-0.294	-0.326	-0.248	-0.438
+1.013	+0.929	+0.914	-0.980	+1.021	-0.398	-0.294	-0.495	-0.186	-0.426
+1.396	-0.269	+0.515	+0.490	+0.116	-0.641	-0.112	-0.574	+1.493	+0.114
-1.093	-0.919	+1.312	-0.490	+0.116	-0.654	-0.721	-0.610	+1.337	-0.180
-1.093	+1.431	-0.282	+0.122	+0.429	-0.641	-0.538	-0.471	-0.062	-0.570
+0.630	+0.788	+0.914	-1.960	+0.429	-0.641	-0.721	-0.625	-0.777	-0.834
+1.013	+0.840	-0.282	-0.980	-0.197	+0.228	-0.294	-0.462	-0.466	-0.570
-0.710	-0.113	+1.312	-1.347	+0.604	-0.501	-0.965	-0.610	-0.478	-0.672
-0.901	+2.097	-0.282	-0.857	+0.429	-0.692	-0.782	-0.628	-0.932	-0.852

Table 3:  
Contents of the "cn1" file. It is the "Meau" file after normalisation (centering and reduction) by column.

The "cn1.CPMA" file, which can also be translated into text, contains the following values (as for all the binary files the structuring in row/columns is fictitious):

24	10	2	10						
7.708	196.3	8.271	325	87.67	6.717	2.083	2.12	5.198	1.55
27.29	18305	0.063	1667	824.9	61.22	2.695	10.96	10.34	2.781

Among the first values one finds the number of rows(24), the number of columns (10) the type of PCA (normalized = 2) and the total inertia (10). The second line gives the means of the 10 variables and the third line gives their variances.

After this "Preparation" stage one just uses the "Computation" command of the "PCA" menu. One indicates the analysis title for which the calculations are to be performed: here it is "cn1" (figure 3.3) and the list of the eigenvalues is displayed:

Num.	Eigenval.	%Iner.	R.Sum	Num.	Eigenval.	%Iner.	R.Sum
EV 1	:0.5745E+01	.5745	.5745	EV 2	:0.1430E+01	.1430	.7175
EV 3	:0.1084E+01	.1084	.8259	EV 4	:0.6761E+00	.0676	.8935
EV 5	:0.5244E+00	.0524	.9459	EV 6	:0.3033E+00	.0303	.9763
EV 7	:0.1462E+00	.0146	.9909	EV 8	:0.5371E-01	.0054	.9963
EV 9	:0.2273E-01	.0023	.9985	EV 10	:0.1454E-01	.0015	1.0000

In this list one can see that the first four eigenvalues total almost 90% of the total inertia (89.35%) and one chooses to calculate the factorial coordinates on the first four factors (figure 3.5). Another method for determining the number of factors to save would consist in drawing the graph of eigenvalues and choosing those which precede a break in the gradient (figure 6.3).

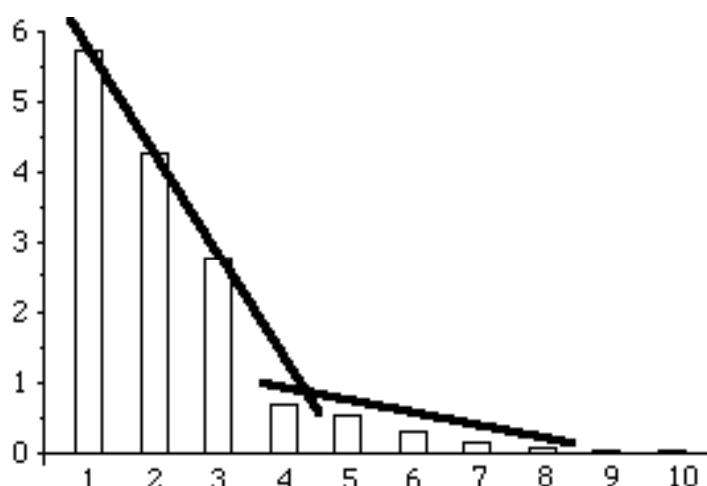


Figure 6.3:

Diagram of the eigenvalues of a PCA showing a method for selecting the number of factors to save. Here one can see a break in the gradient which leads to saving only the first three factors.

The following files are created during this stage (figure 6.4, 6.5 and 6.6):

The "cn1.CPVP" file (table 4) contains the list of eigenvalues (this allows a graphic representation like that of figure 6.3), preceded by the number of factors saved and the total inertia (here 10 since one is dealing with a normalized PCA and there are ten variables).

4  
10  
5.745  
1.43  
1.084  
0.6761  
0.5244  
0.3033  
0.1462  
0.0537  
0.0227  
0.0145

Table 4:

*Contents of the "cn1.CPVP" file. Note that the first two values correspond to the number of factors saved (4) and to the total inertia (10). These values must obviously not be used to draw a graphical representation like the one in figure 6.3.*

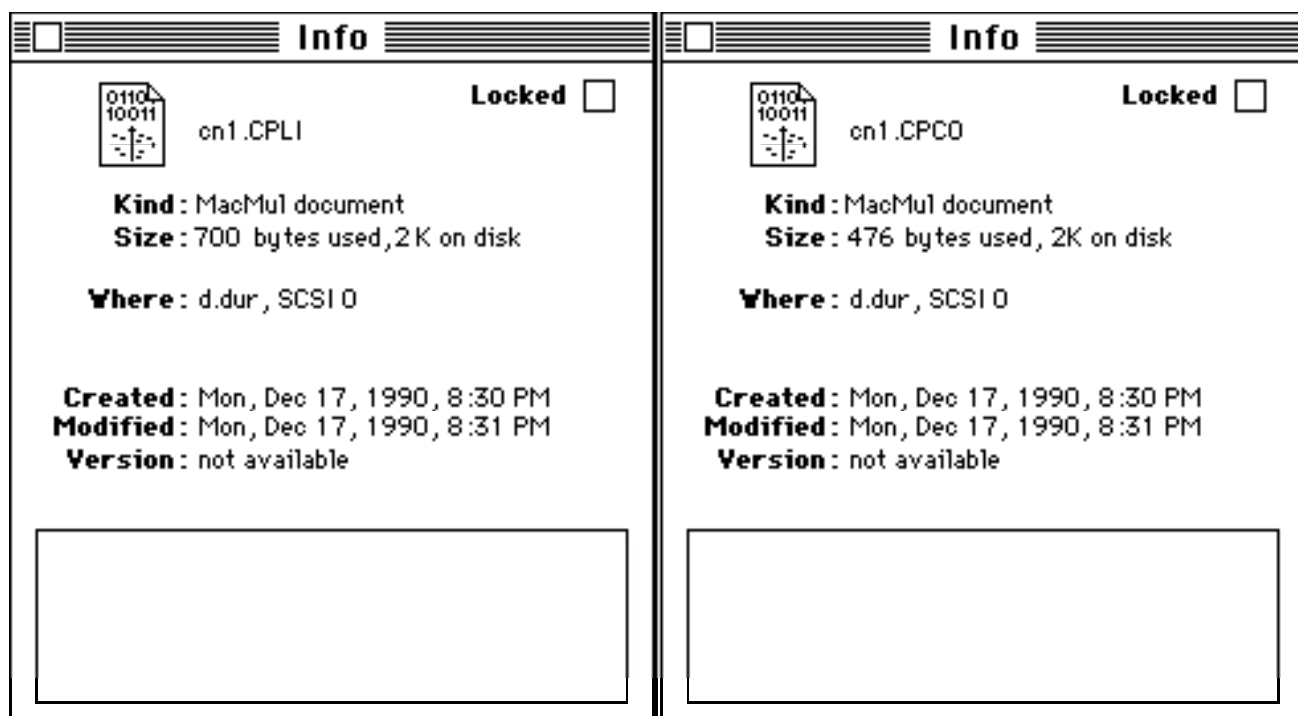


Figure 6.4:

*Information windows of the "cn1.CPLI" and "cn1.CPCO" files, which contain the factorial coordinates of rows and columns.*

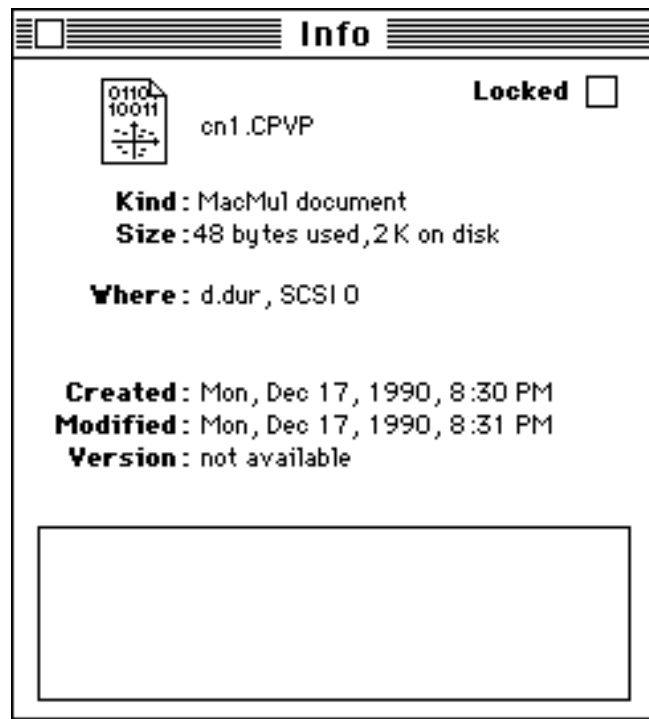


Figure 6.5:  
Information window of the "cn1.CPVP" file.

The contents of the "cn1.CPLI" and "cn1.CPCO" files is the following (tables 5 and 6):

1.571	-.079	1.020	1.177
.757	-.371	1.333	.658
1.408	1.258	-.171	1.434
.742	-.290	-.482	.896
-1.263	-1.135	1.050	.781
-5.389	-1.543	.748	.275
-7.234	-.397	-1.483	1.163
-.407	.045	-1.280	.105
1.723	-.534	.634	.236
-3.721	.652	1.222	-1.197
-1.049	2.927	-.452	-.480
-.679	.113	-1.578	-.446
1.932	-.703	.397	-.290
-2.102	.639	1.293	-1.285
.776	2.511	-.196	-.079
1.049	-.815	-1.827	-1.085
1.724	-.827	.494	-.716
.431	1.194	1.382	-1.016
1.448	1.896	-.098	.866
1.165	-.228	-1.525	-.670
2.332	-1.378	.585	.004
.692	-1.313	.671	-.638
2.210	-.167	-.218	1.010
1.881	-1.455	-1.521	-.702

Table 5:  
Contents of the "cn1.CPLI" file. The 4 factors saved in the analysis are in columns. The table gives, for each row, its factorial coordinate on each of the four factors.



-.134	-.324	.855	-.342
.380	-.515	-.489	-.572
.789	.066	.014	.228
-.869	.265	-.204	-.116
.814	.133	-.199	-.110
-.942	-.145	-.106	.056
-.922	-.186	-.083	.117
-.978	-.085	-.030	-.007
.057	.920	.096	-.332
-.920	.240	-.069	-.166

Table 6:  
Contents of the "cn1.CPCO" file. The 4 factors saved in the analysis are in columns. The table gives, for each column, its factorial coordinate on each of the four factors.

The two last files created are the "cn1.CPCL" and "cn1.CPCC" files (figure 6.6). They are only used for the inertia analysis; they have the following contents (table 7):

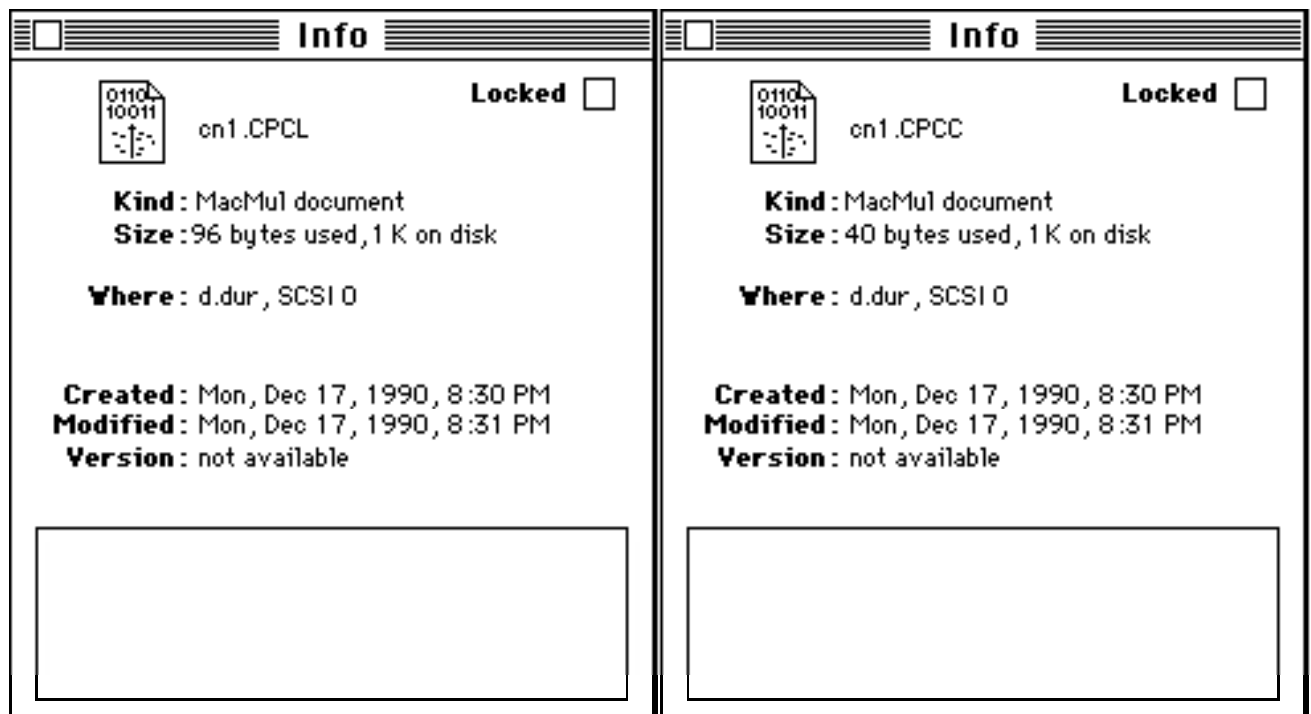


Figure 6.6:  
Information windows of the "cn1.CPCL" and "cn1.CPCC" files, which contain the contributions to the trace of rows and columns.

.0225	.1000
.0167	.1000
.0263	.1000
.0223	.1000
.0355	.1000
.1388	.1000
.2435	.1000
.0097	.1000
.0176	.1000
.0759	.1000
.0446	
.0163	
.0259	
.0346	
.0300	
.0271	
.0226	
.0231	
.0298	
.0199	
.0354	
.0154	
.0277	
.0387	

*Table 7:  
Content of the "cn1.CPCL" file (left column)  
and "cn1.CPCC" file (right column).*

The "cn1.CPLI" and "cn1.CPCO" files are enough to draw the factorial diagrams of rows and columns with GraphMu. On selecting the default options, one immediately obtains figures 6.7 and 6.8. For the rows (figure 6.7) the minimum and maximum of the abscissae are [-7.3, +2.7] and those of the ordinates are [-5, +5].

For the columns (figure 6.8) the minimum and maximum of the abscissae are [-1, +1] and those of the ordinates are [-1, +1] (i.e. the correlation circle; the circle and various legends can easily be added, e.g. with MacDraw).

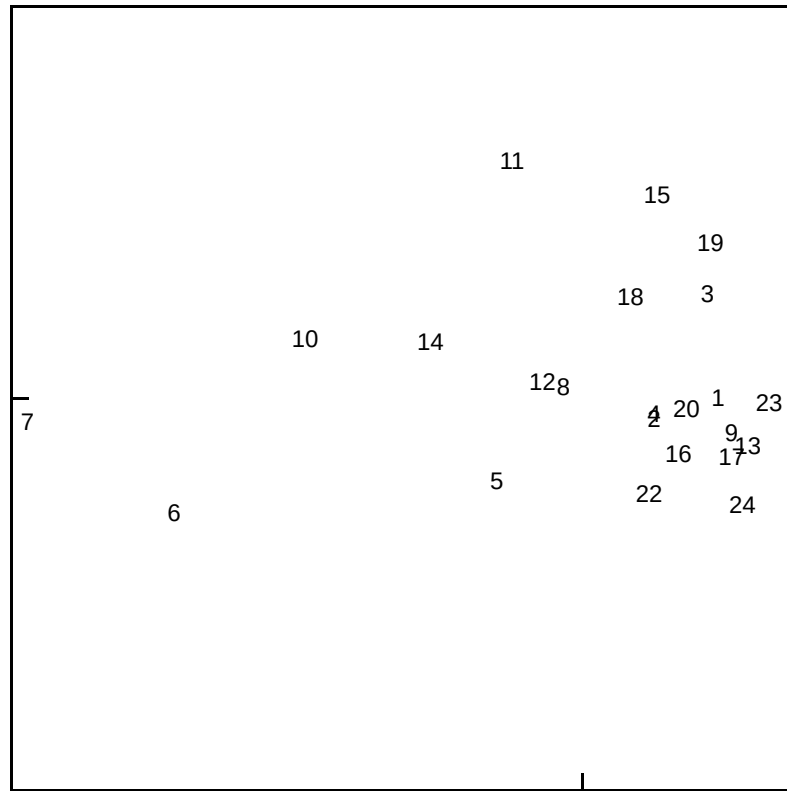


Figure 6.7:  
Factorial diagram (F1 x F2) of the rows, drawn with GraphMu and the "cn1.CPLI" file.

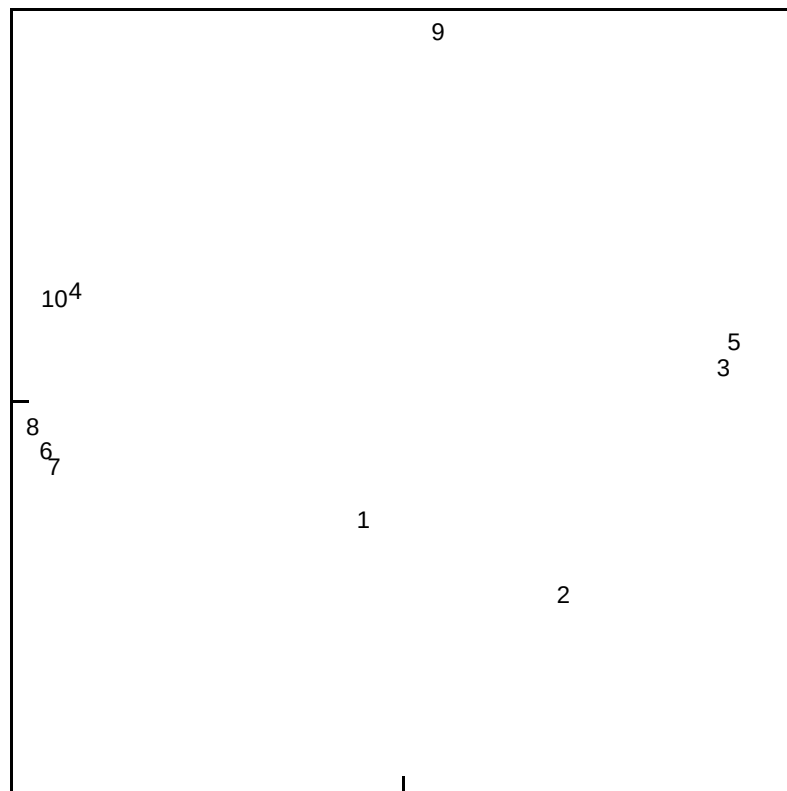


Figure 6.8:  
Factorial diagram (F1 x F2) of the columns, drawn with GraphMu and the "cn1.CPCO" file.

## PCA: analysis on profiles

A small numerical example will be used to show the use of the MacMul "PCA on percentages" command. One sees in table 8 that this example has 3 columns and 6 rows. The 3 columns correspond to 3 granulometric classes and the 6 rows to 6 samples.

0.5	0.3	0.2
0.7	0.2	0.1
0.8	0.1	0.1
0.0	0.9	0.1
0.1	0.2	0.7
0.2	0.2	0.6

Table 8:  
*Example of a table for a PCA on profiles.*

After having run the **centered** PCA of this table ("Preparation" and "Computation" commands) one can use the "PCA on percentages" command. The user gives the analysis title, here "prof" and obtains the "prof.CPPO" and "prof. CPVE" files, the contents of which is shown in tables 9 and 10. The "prof. CPPO" file has 7 rows (the first corresponds to the center of gravity) and the "prof. CPVE" file has 3. Using GraphMu one can draw figure 6.9.

.059	-.020
.192	-.099
.435	-.133
.566	-.080
-.484	-.503
-.233	.390
-.122	.303

Table 9:  
*Contents of "prof.CPPO" file*

.808	-.115
-.504	-.643
-.305	.758

Table 10:  
*Contents of "prof.CPVE" file*

To do this, one must first of all draw a "Characters map" with the "prof.CPPO" file as XY file and a text file containing on the first line a star (\*) and on the following lines the figures 1, 2, 3, ..., 6 as G file. A second map with characters is drawn with the "prof. CPVE" file as XY file and no G file (by default, these are the numbers 1, 2, 3 which are used), to mark the triangle's vertices. A dots curve with

the "prof. CPPO" file as X file **and** G file allows the crosses (+) to be placed corresponding to the exact position of each sample. Finally, a line curve with the "prof.CPVE" file is used to draw the triangle sides.

All these graphs are superimposed by copy/paste and the whole is pasted in the MacDraw II program, which is used for some retouching (closing the triangle, moving characters, ...) and the rotation which allows us to obtain the triangle at the right of figure 6.9.

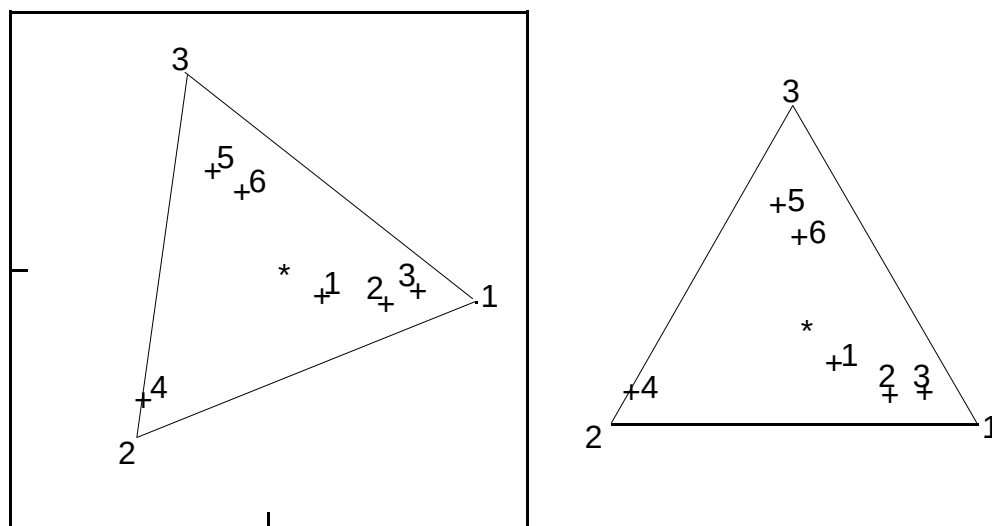


Figure 6.9:  
Figure drawn with GraphMu and the "prof.CPPO" and "prof.CPVE" files. The figure at the right has been rotated using the MacDraw II program.

One can see easily that this representation conforms well with the classical "triangular representation" used in granulometry.

## CA: "double discrimination"

The numerical example used here is taken from the following paper:

Pirot J.Y., Chessel D., et Tamisier A. (1984) Exploitation alimentaire des zones humides de Camargue (delta du Rhône, France) par cinq espèces de canards de surface hivernants: modélisation spatio-temporelle. *Terre et Vie*, 39, 167-192.

It involves the results of samples by boring in 10 types of stations, supplying 30 plant categories (table 11). The research into the division of resources between the species of ducks living in the Camargue, either overwintering or migrating, requires a good knowledge of the descriptive capacity of these samples. This table is recorded in the "carot.txt" file and transformed into binary in the "carot" file

	a	b	c	d	e	f	g	h	i	j
1	6	0	0	0	0	0	0	0	0	0
2	7	6	0	0	0	0	0	0	0	0
3	7	6	0	0	0	0	0	0	0	0
4	9	8	0	0	0	0	0	0	0	0
5	10	8	2	0	0	0	0	0	0	0
6	6	4	2	6	0	0	0	0	0	0
7	1	0	0	0	0	0	1	0	0	0
8	0	4	6	0	0	0	0	0	0	0
9	9	9	10	9	11	6	2	4	5	0
10	9	5	8	7	0	5	8	2	4	3
11	0	2	5	6	5	6	0	3	0	0
12	0	0	6	6	0	2	0	1	0	0
13	0	0	6	9	7	2	0	0	0	0
14	0	2	5	7	10	4	2	3	2	0
15	0	0	0	3	0	0	0	0	0	0
16	1	1	9	2	3	4	0	5	3	0
17	0	0	2	9	12	3	1	3	2	0
18	0	0	3	0	0	0	0	0	2	0
19	0	0	9	9	9	8	8	7	4	0
20	0	0	0	11	7	6	5	0	4	0
21	0	0	0	4	1	4	0	4	1	0
22	0	0	4	3	1	6	3	8	6	6
23	0	0	0	0	0	0	3	0	2	0
24	0	0	0	0	0	0	0	3	0	0
25	0	0	0	0	0	1	4	3	3	3
26	0	0	0	0	0	0	1	2	6	0
27	0	0	0	3	0	0	6	0	5	9
28	0	0	0	0	0	0	5	7	4	4
29	0	0	0	0	0	0	0	0	3	0
30	0	0	0	0	0	0	0	3	6	3

Table 11:

The 10 columns correspond to the 10 following types of samples:

*a. Cultivated media (sample 2), b: Cultivated media (sample 1), c: Brackish marsh, d: Area with Cyperaceae (sample 2), e: Area with Cyperaceae (sample 1), f: Grass (sample 2), g: Grass (sample 1), h: "Sansouires", i: Saltwort area, j: Salt marshes. The 30 rows correspond to the 30 plant categories of table 12.*

The CA of the "carot" file gives the row and column coordinates ("Preparation" and "Computation" commands, analysis title = "fcarot"). One then use the "Transpose file" command of the "File" menu to transpose the "carot" file, then the "Double discrimination" command, using just the first two factors. The files created are the "fcarot.FCCA" (coding of the cells of the contingency table, 300 rows and 3 columns), "fcarot.FCVL" (row variances in the form of an MVC file, 30 rows and 5 columns) and "fcarot.FCVC" (columns variances in the form of an MVC file, 10 rows and 5 columns). The contents of the "fcarot.FCVL" and "fcarot.FCVC" files is given in tables 13 and 14.

1: <i>Oriza sativa</i>	16: <i>Eleocharis palustris</i>
2: <i>Echinocloa sp.</i>	17: <i>Chara sp. (large)</i>
3: <i>Naias sp. (large)</i>	18: <i>Atriplex hastata</i>
4: <i>Naias sp. (small)</i>	19: <i>Ranunculus sp.</i>
5: <i>Scirpus mucronatus</i>	20: <i>Potamogeton pectinatus</i>
6: <i>Potamogeton pusillus</i>	21: <i>Suaeda fruticosa</i>
7: <i>Obione portulacoides</i>	22: <i>Arthrocnemum glaucum</i>
8: <i>Paspalum distichum</i>	23: <i>Suaeda sp. (large)</i>
9: <i>Chara sp. (small)</i>	24: <i>Trifolium sp.</i>
10: <i>Scirpus maritimus</i>	25: <i>Suaeda sp. (small)</i>
11: <i>Zannichelia sp.</i>	26: <i>Salicornia herbacea</i>
12: <i>Myriophyllum spicatum</i>	27: <i>Ruppia maritima</i>
13: <i>Scirpus lacustris</i>	28: <i>Salicornia fruticosa</i>
14: <i>Scirpus littoralis</i>	29: <i>Kochia hirsuta</i>
15: <i>Potamogeton sp. (remains)</i>	30: <i>Ruppia cirrhosa</i>

Table 12:  
List of the 30 plant categories of table 11.

2.494	.862	.000	.000	.000
2.329	.610	.006	.008	.011
2.329	.610	.006	.008	.011
2.326	.605	.006	.008	.011
2.098	.446	.105	.071	.051
1.184	-.213	.338	.220	.145
.853	1.149	.527	-.081	.012
.830	-.542	.223	.129	.074
.307	-.371	.316	.098	.207
.239	.320	.356	-.016	.282
-.227	-.961	.098	.027	.085
-.284	-1.072	.011	-.011	.034
-.325	-1.385	.006	.005	.012
-.323	-.798	.084	-.004	.189
-.352	-1.342	.000	.000	.000
-.255	-.446	.122	-.013	.186
-.494	-.998	.009	-.031	.176
-.407	.028	.039	-.108	.296
-.521	-.430	.016	-.048	.238
-.548	-.576	.008	-.044	.272
-.598	-.489	.010	-.039	.167
-.735	.664	.024	-.092	.423
-.855	1.559	.001	-.002	.003
-.866	.624	.000	.000	.000
-.915	1.597	.007	-.037	.220
-.917	1.460	.001	-.003	.031
-.939	1.946	.016	-.077	.366
-.937	1.625	.005	-.024	.160
-.955	1.743	.000	.000	.000
-1.001	1.901	.004	-.024	.159

Table 13:  
"fcarot.FCVL" file.

1.979	.548	.269	.098	.064
1.695	.201	.305	.149	.099
-.033	-.708	.102	.027	.108
-.280	-.852	.072	.005	.170
-.378	-1.136	.027	.008	.050
-.409	-.623	.035	-.008	.132
-.626	.912	.070	-.079	.369
-.688	.396	.043	-.087	.355
-.758	1.107	.056	-.107	.392
-.974	2.219	.041	-.059	.140

Table 14:  
"fcarot.FCVC" file.

In the "fcarot.FCCA" file only the rows corresponding to non-zero entries are saved: this operation can be performed after the "fcarot.FCCA" file has been transformed into a text file (which will later be retransformed into binary), either "by hand" or automatically for the large tables with a spreadsheet, e.g. Excel or StatView.

2.494	.862	6.000
2.402	.708	7.000
2.244	.496	6.000
2.402	.708	7.000
2.244	.496	6.000
2.400	.705	9.000
2.242	.493	8.000
2.273	.608	10.000
2.115	.396	8.000
1.151	-.160	2.000
1.764	.204	6.000
1.605	-.008	4.000
.642	-.564	2.000
.504	-.652	6.000
1.579	1.038	1.000
etc.		

Table 15:  
The first 15 rows of the "fcarot.FCCA" file after suppression of the rows corresponding to a zero entry. The first two columns correspond to the first two factors and the third contains the entry of the considered cell.

Figure 6.10 was drawn with GraphMu in three stages and in windows of size 700 x 1050:

- 1 - "Ellipses" command of the "Graphics" menu with the "fcarot.FCVC" file to draw the ellipses corresponding to the columns (samples, at the top),
- 2 - "Ellipses" command of the "Graphics" menu with the "fcarot.FCVL" file to draw the ellipses corresponding to the rows (species, at the bottom),
- 3 - "Dot curves" command with the "fcarot.FCCA" file to draw the crosses symbolizing each cell of the table.

The image obtained at stage 3 is superimposed over those obtained in stages 1 and 2 by Copy/Paste in MacDraw II. The whole is reduced so as to obtain a figure which



fits into a page; the size of the crosses ("+" sign in Monaco font) is set to 6 points and the thickness of the lines is set at 0.1 points. Characters A, B, C, D and figures 1 to 7 are then added, in 18-point Times.

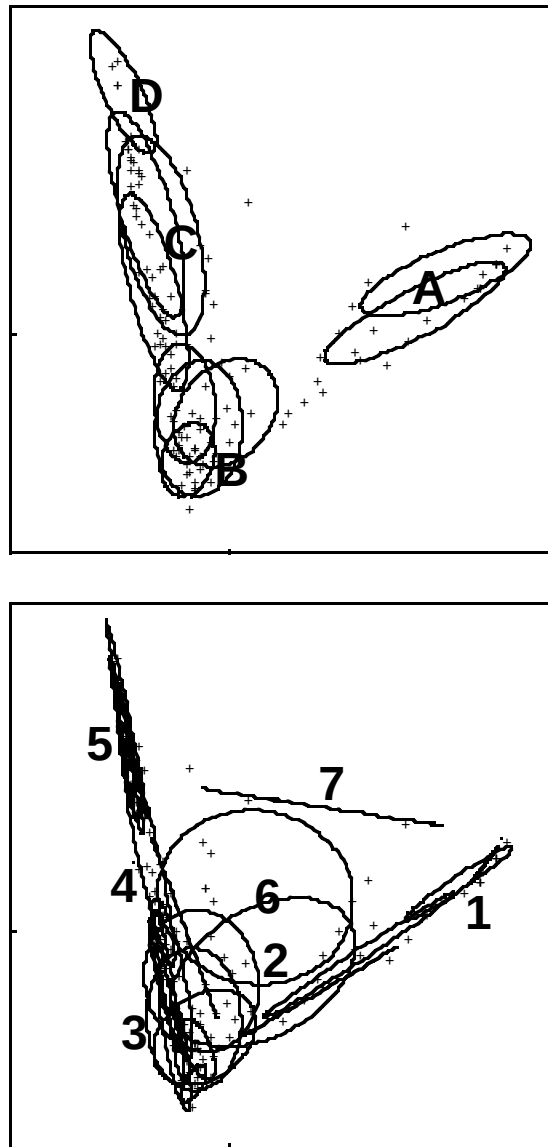


Figure 6.10:

Figure drawn with GraphMu and the "fcarot.FCCA" and "fcarot.FCVC" files (at the top) or "fcarot.FCVL" (at the bottom). This figures was drawn in a window size 700 x 1050, then reduced with MacDraw II (character size = 6 points, line thickness = 0.1 points).

The authors give the following interpretation: each column (sampling point) is described by its position and its factorial diversity, each row (taxon) is described by its position and factorial amplitude, in both cases by inertia ellipses. The biomasses present in the soil allow clear identification of four types of medium (A: cultivated, B: fresh, C: brackish, D: salty) and seven types of species defined as a function of the position **and amplitude**. There is clear ambiguity between the taxons of groups 2, 6 and 7: the occurrence of such a plant species in the stomach contents of a duck will impede identification of the place of feeding.

## MCA: "Correlation ratios"

The data set used here comes from the work of D. Pontier (1984): Contribution a la biologie et la Génétique de chats domestiques (*Felis catus*). Thèse de 3° cycle. Université Lyon 1.

Table 16 contains on the left 3 qualitative variables (in column) measured on 26 groups of cats (rows) and on the right the number of cats for each group. The variables and their categories are described below (total of 12 categories):

Column 1	age:	1 = 1 year 2 = 2 or 3 years 3 = 4 or 5 years 4 = 6 or 7 years 5 = 8 years or more
Column 2	fertility:	1 = 1 or 2 kittens/year 2 = 3 to 6 kittens/year 3 = 7 or 8 kittens/year 4 = 9 to 12 kittens/year 5 = 13 or 14 kittens/year
Column 3	Number of litters/year	1 = 1 litter 2 = 2 litters

1 1 1	7
2 1 1	3
5 1 1	2
1 2 1	18
2 2 1	14
3 2 1	3
4 2 1	3
5 2 1	3
1 2 2	5
2 2 2	10
3 2 2	1
4 2 2	5
5 2 2	3
3 3 1	1
1 3 2	1
2 3 2	14
3 3 2	7
4 3 2	5
5 3 2	5
1 4 2	1
2 4 2	4
3 4 2	6
4 4 2	4
5 4 2	5
2 5 2	2
4 5 2	2

Table 16:

*The three columns at the left are the categories of the three qualitative variables and the column at the right is the number of cats presenting one of the 26 category combinations. As an example, 18 one-year old cats were observed, which had had 3 to 6 kittens, with one litter per year.*

The three qualitative variables were put in a text file ("chat.txt") and the cat numbers were typed in another file ("chateff.txt"). These files were transformed into binary ("chat" and "chateff" files) and the MCA was run with MacMul ("Preparations" and "Computation" commands, analysis title = "cma"). The non uniform row weights option was chosen during preparation (figure 5.1), and the "chateff" file was indicated as row weighting file (figure 5.2). The "Computation" command then supplies the factorial coordinates files "cma.MCLI" and "cma.MCCO" with 6 factors saved (tables 17 and 18).

-1.527	-.584	-.352	.743	1.162	-.013
-1.003	-.194	.243	1.094	.809	.621
-.887	-1.070	-.512	1.712	.245	-.090
-1.144	.022	-.118	-.481	-.104	-.205
-.620	.411	.477	-.130	-.457	.429
-.306	-.456	.246	-1.048	.361	.196
-.376	.686	-.686	-.487	-.156	-.731
-.504	-.465	-.278	.487	-1.021	-.282
-.436	.071	-.129	-.412	-.145	-.248
.088	.461	.467	-.061	-.498	.387
.402	-.406	.236	-.979	.320	.153
.332	.735	-.696	-.418	-.196	-.773
.204	-.415	-.289	.556	-1.062	-.324
.297	-.633	.740	-.572	.953	-.272
.167	-.107	.365	.064	.447	-.715
.691	.283	.960	.415	.094	-.081
1.005	-.584	.729	-.503	.912	-.314
.935	.558	-.202	.058	.396	-1.241
.807	-.593	.205	1.032	-.470	-.792
.214	-.616	-.791	-.366	-.003	.432
.739	-.226	-.195	-.015	-.356	1.066
1.053	-1.093	-.427	-.932	.461	.833
.982	.048	-1.358	-.371	-.055	-.094
.854	-1.102	-.951	.603	-.920	.355
.717	2.068	-.645	.945	.841	1.578
.961	2.342	-1.808	.588	1.143	.418

Table 17:  
Factorial coordinates of the rows ("cma.MCLI" file).

-1.235	-.197	-.287	-.311	.328	-.309
.083	.559	.811	.307	-.275	.714
.871	-1.122	.385	-1.307	1.122	.337
.695	1.091	-1.335	-.320	.240	-1.156
.374	-1.139	-.583	1.394	-1.239	-.432
-1.540	-.878	-.374	1.691	1.616	.247
-.578	.296	.058	-.464	-.547	-.063
.937	-.049	.969	.373	.465	-.817
1.056	-1.037	-1.164	-.383	-.305	1.032
1.002	3.412	-1.994	1.307	1.741	1.857
-1.062	-.057	.012	-.073	.042	.041
.717	.039	-.008	.049	-.028	-.027

Table 18:  
Factorial coordinates of categories ("cma.MCCO" file).

The "Correlation ratios" command allows us to obtain the "cma.MCRC" file, the contents of which is shown below (six factors analysed). One finds the number of cats belonging to each category of three variables, with the corresponding frequencies. The correlation ratios table shows that the first factor is quite closely linked to the three variables (especially to variables 2 and 3), but that variable number 3 is practically uncorrelated with factors 2 and 3.

MCA: Correlation ratios.

Title of analysis:

cm2 26 rows 3 variables 12 categories

Description of categories:

Variable number 1 has 5 categories:

1   Cat.: 1	Ind.: 32	Freq.: .2388
2   Cat.: 2	Ind.: 47	Freq.: .3507
3   Cat.: 3	Ind.: 18	Freq.: .1343
4   Cat.: 4	Ind.: 19	Freq.: .1418
5   Cat.: 5	Ind.: 18	Freq.: .1343

Variable number 2 has 5 categories:

6   Cat.: 1	Ind.: 12	Freq.: .0896
7   Cat.: 2	Ind.: 65	Freq.: .4851
8   Cat.: 3	Ind.: 33	Freq.: .2463
9   Cat.: 4	Ind.: 20	Freq.: .1493
10   Cat.: 5	Ind.: 4	Freq.: .0299

Variable number 3 has 2 categories:

11   Cat.: 1	Ind.: 54	Freq.: .4030
12   Cat.: 2	Ind.: 80	Freq.: .5970

Variables/Factors correlation ratios.

Factors are in columns.

VAR 1	.5558	.6310	.5688	.5615	.4357	.4315
VAR 2	.7869	.6200	.5662	.4678	.5368	.4336
VAR 3	.7608	.0022	.0001	.0036	.0012	.0011

The groups of cats and the categories can be described thanks to the factorial diagrams of rows and columns. Here we prefer to show an example of use of the "Gauss curves" command of GraphMu. The aim is to obtain a graphic representation of the categories of the three qualitative variables which **take account of the variability** of the numbers of cats belonging to these categories in the 26 groups.

To make this representation with GraphMu, one must have an "MVC" file which will contain as rows the twelve categories of the three variables and as columns the means, variances and covariances of the factorial coordinates of individuals belonging to the categories. The "MVC File" command of the GraphMu "File" menu allows this file to be created from the row factorial coordinates file "cma.MCLI") and from the "complete disjunctive table" associated with the three variables. To obtain this complete disjunctive table from the "chat" file one can use the "O1DI File" command of the GraphMu "File" menu. This command gives the "chat.O1DI" file of table 19.

```

100001000010
010001000010
000011000010
100000100010
010000100010
001000100010
000100100010
000010100010
100000100001
010000100001
001000100001
000100100001
000010100001
001000010010
100000010001
010000010001
001000010001
000100010001
000010010001
100000001001
010000001001
001000001001
000100001001
000010001001
010000000101
000100000101

```

Table 19:

Contents of the "chat.O1DI" file: complete disjunctive table associated with table 16. Each column represents one category of one variable and on each row one finds a 1 or a 0 depending on whether the corresponding cat (or group of cats) belongs to this category or not.

<b>Coordinates file:</b>	<input type="text" value="cma.MCLI"/>	<b>Initial file (AFC) or O1DI (ACM):</b>	<input type="text" value="chat.O1DI"/>
<b>Columns:</b>	<input type="text" value="6"/>	<b>Rows:</b>	<input type="text" value="26"/>
<b>Var. 1:</b>	<input type="text" value="1"/>	<b>Columns:</b>	<input type="text" value="12"/>
<b>Var. 2:</b>	<input type="text" value="2"/>		
<b>Output file (MVC):</b>		<input type="text" value="chatMVC"/>	
<input type="button" value="Cancel"/>		<input type="button" value="OK"/>	

Figure 6.11:

Main dialog window of the "MVC File" command of GraphMu. The MVC file obtained ("chatMVC") is a binary file which can be used directly to draw ellipses or Gauss curves (figure 6.12).

The "MVC File" command can then be used (figure 6.11) and supplies the "chatMVC" file of table 20. The GraphMu "Gauss curves" command is then used to draw figure 6.12.

M1	M2	V1	C12	V2
-.5454	-.2428	.4836	.0049	.0886
.1022	.4670	.4787	.2010	.5863
.4902	-.6344	.2520	-.0789	.0594
.5667	.8738	.2815	.0658	.5986
.0950	-.7289	.4836	.0049	.0886
-1.1388	-.6162	.0776	-.0231	.1283
-.2360	.0644	.2137	-.0006	.2117
.6503	-.1792	.0986	.0243	.2171
.7684	-.5980	.0884	-.0094	.2111
.8390	2.2051	.0148	.0167	.0188
-.6743	-.2537	.2557	.0160	.2676
.5714	.0838	.1666	.0127	.8633

Table 20:

*Contents of the "chatMVC" file. This file is a binary file of "MVC" type with two variables (i.e. the first two factors of the MCA of the "chat" file).*

*The twelve rows correspond to the twelve categories of the three qualitative variables. The first two columns are the means of the factorial coordinates of the groups of cats belonging to the category considered. Columns 3 and 5 contain the variances and column 4 contains the covariance.*

Interpretation of this figure allows analysis of the relations between the categories of the same variable or between categories of the 3 variables. Refer to the work of the Pontier (1984, op. cit.) for a discussion in terms of age-fertility relations.

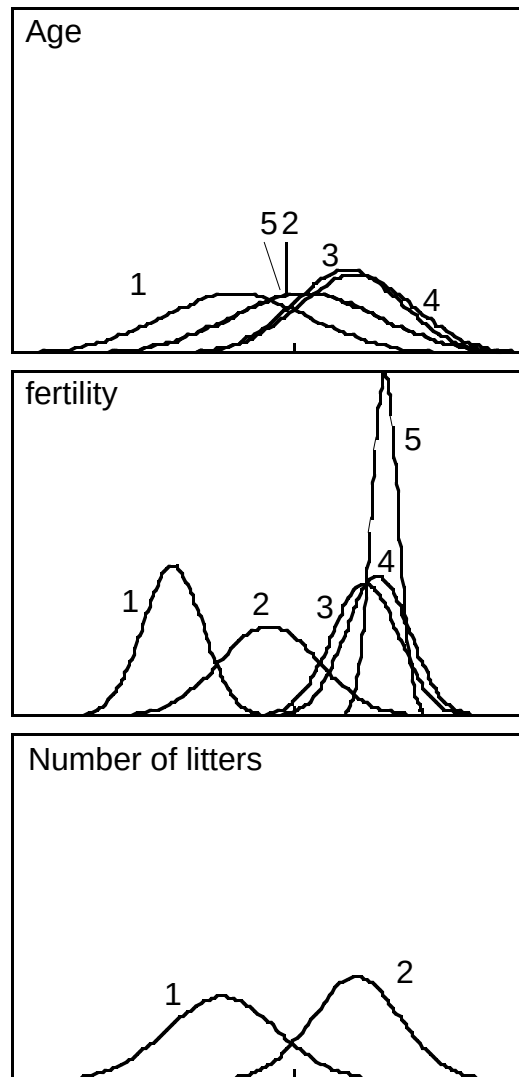


Figure 6.12:  
*Representation by Gauss curves of the twelve categories of the three qualitative variables.*