ANDREWS PLOT

PURPOSE

Generates an Andrews plot.

DESCRIPTION

An Andrews plot is a graphical data analysis technique for plotting multivariate data. An Andrews curve applies the following transformation to a set of data:

$$F_i(t) = X1_i/SQRT(2) + X2_i*SIN(t) + X3_i*COS(t) + X4_i*SIN(2t) + X5_i*COS(2t) + ...$$

where t goes from $-\pi$ to π and X1, X2, etc. are the columns (i.e., variables) of data. One Andrews curve is generated for each row of data. As usual, the LINE, LINE COLOR, and LINE THICKNESS commands can be used to control the attributes of the curves. Andrews curves are used to distinguish which observations (i.e., rows) are most alike.

SYNTAX

ANDREWS PLOT <y1> <y2> ... <yk> <SUBSET/EXCEPT/FOR qualification>
where <y1> through <yk> are the response variables;
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

EXAMPLES

ANDREWS PLOT Y1 Y2 Y3 Y4 Y5
ANDREWS PLOT Y1 Y2 Y3 Y4 Y5 SUBSET TAG > 2

NOTE 1

The increment for t in the transformation can be set with the ANDREWS INCREMENT command. It defaults to 0.1.

NOTE 2

Andrews curves are order dependent. The first few variables tend to dominate, so it is a good idea to put the most important variables first. Some analysts recommend running a principle components analysis first and generating Andrews curves for the principle components.

A related plot which is not order dependent is the parallel coordinates plot. The plot is divided into a series of parallel axes (one for each variable). An observation is then generated by plotting its value on each axis and connecting them between axes with a line. Although DATAPLOT does not generate this plot directly, a macro to generate one is demonstrated in program example 2.

With both of these plots, individual cases can be difficult to follow. It can sometimes help to draw the plot for all cases and then draw the plot over various subsets of interest (with the subsets having a limited number of cases).

NOTE 3

Up to 20 variables can be used.

NOTE 4

The TO syntax is allowed on this command. For example:

ANDREWS PLOT Y1 TO Y10

DEFAULT

None

SYNONYMS

None

RELATED COMMANDS

LINES	=	Sets type for plot lines.
LINE COLOR	=	Sets color for plot lines.
LINE THICKNESS	=	Sets thickness for plot lines.
PLOT	=	Generates a data or function plot.
STAR PLOT	=	Generates a star plot.
PROFILE PLOT	=	Generates a profile plot.
ANDREWS INCREMENT	=	Specify the x axis increment when generating Andrews curves.

REFERENCE

"Graphical Exploratory Data Analysis," du Toit, Steyn, and Stumpf, Springer-Verlang, 1986.

"Hyperdimensional Data Analysis Using Parallel Coordinates," E. J. Wegman, Journal of the American Statistical Association, 85, 664-675.

APPLICATION

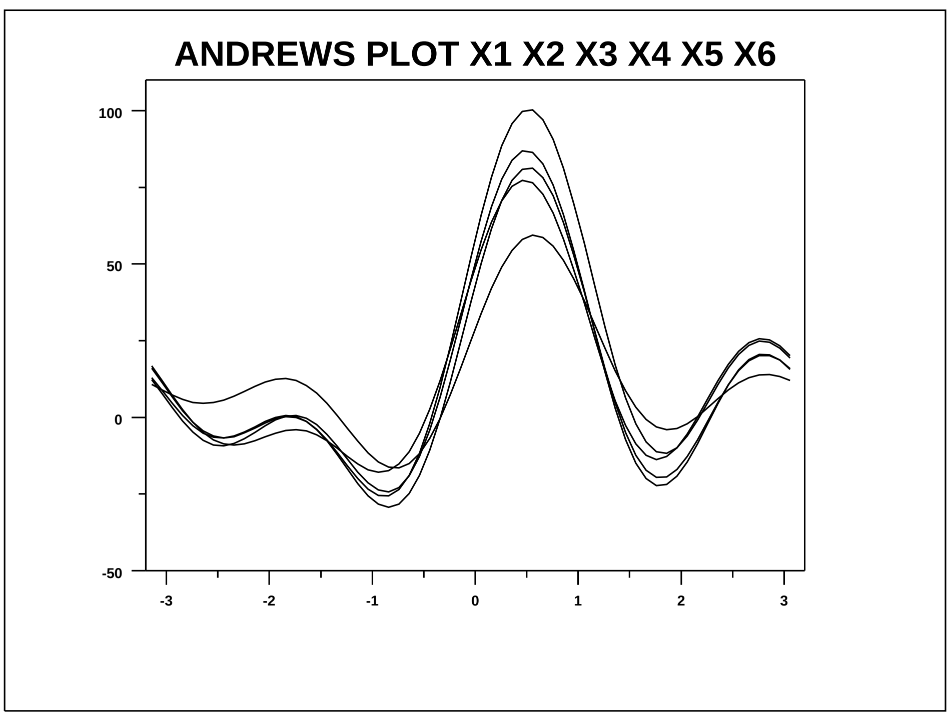
Multivariate Analysis

IMPLEMENTATION DATE

92/12

PROGRAM 1

```
ROW LIMITS 26 50
COLUMN LIMITS 20 132
READ AUTO79.DAT Y1 TO Y9
MULTIPLOT 2 3; MULTIPLOT CORNER COORDINATES 0 0 100 100
TITLE AUTOMATIC; TITLE SIZE 3.0
TIC LABEL SIZE 3
XLIMITS -3 3; XTIC OFFSET 0.2 0.2; MAJOR XTIC MARK NUMBER 7
YLIMITS 0 15000; YTIC OFFSET 1000 1000; MAJOR YTIC MARK NUMBER 18
ANDREWS PLOT Y1 TO Y9
LINES SOLID DASH DOT SOLID DASH DOT
LINE THICKNESS 0.1 0.1 0.1 0.3 0.3 0.3
ANDREWS PLOT Y1 TO Y9 FOR I = 1 1 5
ANDREWS PLOT Y1 TO Y9 FOR I = 6 1 10
ANDREWS PLOT Y1 TO Y9 FOR I = 11 1 15
ANDREWS PLOT Y1 TO Y9 FOR I = 16 1 20
ANDREWS PLOT Y1 TO Y9 FOR I = 21 1 25
END OF MULTIPLOT
```



PROGRAM 2

```
DIMENSION 20 COLUMNS
LET P = 9
ROW LIMITS 26 50; COLUMN LIMITS 20 132
READ AUTO79.DAT X1 TO X^P
.
LET N = SIZE X1; LET 2N = 2*N; LET TEMP = N + 1
LET Y = 0 FOR I = 1 1 N; LET Y = 1 FOR I = TEMP 1 2N
LET TAG = SEQUENCE 1 1 N FOR I = 1 1 2N
.
LOOP FOR K = 1 1 P
    LET M = MEAN X^K; LET SD = STANDARD DEVIATION X^K
     LET X^K = (X^K - M)/SD
END OF LOOP
.
LET TEMP = P - 1
MULTIPLOT TEMP 1; MULTIPLOT CORNER COORDINATES 5 5 95 95
FRAME CORNER COORDINATES 5 0 95 100; YFRAME OFF
TIC LABELS OFF; TIC MARKS OFF
YLIMITS 0 1; XLIMITS -3.5 3.5
LEGEND 1 COORDINATES 4.5 98; LEGEND JUSTIFICATION RIGHT; LEGEND SIZE 12
.
LOOP FOR K = P -1 2
    LET A = K - 1; LET B = K; LET X = X^A
    EXTEND X X^B; LEGEND 1 X^B
    PLOT Y X TAG
END OF LOOP
JUSTIFICATION RIGHT; MOVE 4.5 1.5; HEIGHT 12; TEXT X1
END OF MULTIPLOT
HEIGHT; JUSTIFICATION CENTER; MOVE 50 97; TEXT PARALLEL COORDINATES PLOT
```