... SMOOTH

PURPOSE

Performs

1. least squares smoothing;

2. moving average (or some alternate statistic) type smoothing;

3. robust smoothing.

DESCRIPTION

These are data analysis techniques for producing a "smooth" set of values from a time series which has been contaminated with noise. For least squares smoothing, the analyst can specify the degree (default = 1) and width (default = 3) of the smoothing function. Any smoothing function from degree 0 to degree 5 may be used (moving average smoothing is mathematically equivalent to both degree 0 and degree 1 smoothing). To specify the desired degree, the POLYNOMIAL DEGREE command should be used prior to the SMOOTH command. Alternatively (and more commonly) the desired degree can be specified as the first word of the SMOOTH command. Thus

> POLYNOMIAL DEGREE 3
> SMOOTH Y

can be shortened to

> CUBIC SMOOTH Y

To specify the width of the least squares smoothing function, the FILTER WIDTH command should be used prior to the SMOOTH command. The filter width must be odd.

For robust smoothing, the 3RSR smoother (developed by John Tukey) is used. For the 3RSR smoother, the filter width is always 3 (and so the FILTER WIDTH command is not needed) and the filter degree is not used. The 3RSR smoother involves 4 steps:

1. computes 3-term moving medians;

2. computes 3-term medians of the 3-term medians repeatedly until no change;

3. applies a "splitting" process;

4. repeats the splitting process until no change.

See the Tukey book listed in the REFERENCE section for complete details of the robust smoothing.

For moving average (or some other statistic) type smoothing, the filter width controls how many values are used in the computation of the statistic. As with least squares smoothing, the filter width should be odd. The smoothing degree is ignored for this type of smoothing (it is implied by the choice of statistic).

For moving average and least squares smoothing, the choice of filter width controls how "smooth" the smoothed line is. In general, shorter filter widths will follow the individual data points more. For least squares smoothing, the degree also affects the smoothness. In general, higher degrees will follow the individual data points more.

SYNTAX 1

<d> SMOOTH <y>   <SUBSET/EXCEPT/FOR qualification>

where <d> is the specification of the desired degree:

> CONSTANT or ZERO-DEGREE
> LINEAR or FIRST-DEGREE
> QUADRATIC or SECOND-DEGREE
> CUBIC or THIRD-DEGREE
> QUARTIC or FOURTH-DEGREE
> QUINTIC or FIFTH-DEGREE

<y> is the variable under analysis;

and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax generates a least squares smooth.

SYNTAX 2

ROBUST SMOOTH <y>   <SUBSET/EXCEPT/FOR qualification>

where <y> is the variable under analysis;

and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax generates the Tukey smoother.

SYNTAX 3

<stat> SMOOTH <y>   <SUBSET/EXCEPT/FOR qualification>

where <stat> is the specification of the type of smoothing:

        MOVING AVERAGE
        MOVING MEDIAN
        MOVING MIDMEAN
        MOVING MIDRANGE
        MOVING UPPER QUARTILE
        MOVING LOWER QUARTILE
        MOVING MAXIMUM
        MOVING MINIMUM
        MOVING TRIANGLE
        MOVING HAMMING;

    <y> is the variable under analysis;

and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax generates a "moving statistic." A MOVING AVERAGE is equivalent to a linear least squares fit and MOVING MEDIAN is the first step in the Tukey smoother (i.e., ROBUST SMOOTH).

EXAMPLES

LINEAR SMOOTH Y
QUADRATIC SMOOTH TEMP
CUBIC SMOOTH V
ROBUST SMOOTH CONC
MOVING MEDIAN SMOOTH Y

NOTE 1

The smoothed values are stored in the internal variable PRED. The internal variable RES contains the difference between the observed values (i.e., the raw data) and the smoothed values.

NOTE 2

See the sources under the REFERENCE section for the mathematical details of the various type of smoothing operations.

NOTE 3

The least squares SMOOTH command assumes equispaced data. The LOWESS SMOOTH command performs locally weighted least squares smoothing and can be used for non-equispaced data.

DEFAULT

For least squares smoothing, the default is degree 1 and filter width 3 (this is equivalent to a moving average smoothing of width 3).

SYNONYMS

MOVING AVERAGE SMOOTH is identical to LINEAR SMOOTH.

RELATED COMMANDS

POLYNOMIAL DEGREE	=	Sets the degree for least squares smoothing.
FILTER WIDTH	=	Sets the filter width for least squares smoothing.
LOWESS SMOOTH	=	Carries out a locally weighted least squares smoothing.
FIT	=	Carries out a least squares fit.

REFERENCES

"Introduction to Numerical Analysis," 2nd. Ed., F. B. Hildebrand, McGraw-Hill, 1973.

"A First Course in Numerical Analysis," Anthony Ralston, McGraw-Hill, 1965, (pp. 250-254).

"Smoothing and Differentiation of Data by Simplified Least Squares Procedure," A. Savitsky M. J. E. Golay, Analytical Chemistry, July, 1964, (pp. 1627-1639).

"Exploratory Data Analysis," John Tukey, Addison-Wesley, 1977.

APPLICATIONS
      Smoothing

IMPLEMENTATION DATE
      Pre-1987

PROGRAM
      SKIP 25
      READ MAVRO.DAT Y
      LET N = SIZE Y
      LET X = SEQUENCE 1 1 N
      .
      MULTIPLOT 2 2; MULTIPLOT CORNER COORDINATES 0 0 100 100
      CHARACTER CIRCLE BLANK
      LINES BLANK SOLID
      CHARACTER SIZE 1.5
      SMOOTH Y
      TITLE DEFAULT SMOOTH
      PLOT Y PRED VS X
      TITLE LEAST SQUARES CUBIC SMOOTH
      CUBIC SMOOTH Y
      PLOT Y PRED VS X
      ROBUST SMOOTH Y
      TITLE ROBUST SMOOTH
      PLOT Y PRED VS X
      MOVING MEDIAN SMOOTH Y
      TITLE MOVING MEDIAN SMOOTH
      PLOT Y PRED VS X
      END OF MULTIPLOT
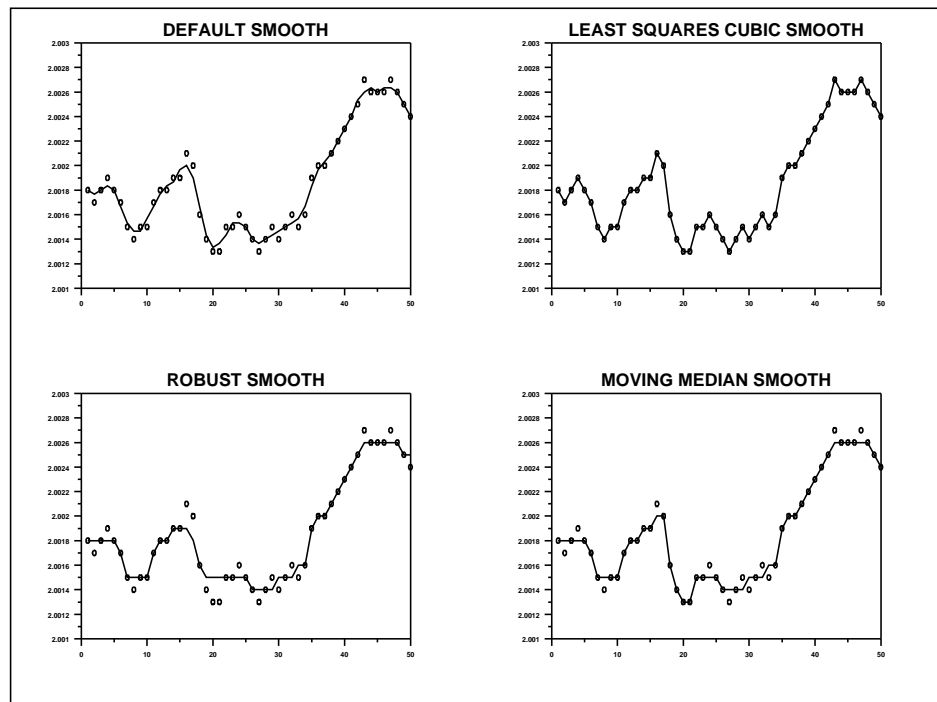
The following output is generated.

```
SMOOTHING

      SMOOTHING FUNCTION--LEAST SQUARES
      NUMBER OF OBSERVATIONS        =        50


**********************************************************************
*                                 *   RESIDUAL    *   AVERAGE      *
*                                 *   STANDARD    *   ABSOLUTE     *
*                                 *   DEVIATION   *   RESIDUAL     *
*                                 * (DIVISOR=N-1) *  (DIVISOR=N)   *
**********************************************************************
* NO SMOOTHING                    *     0.0004291 *      0.0003570 *
* (RAW DATA)                      *               *               *
*                                 *               *               *
* MOVING AVERAGE SMOOTHING        *     0.0000499 *      0.0000387 *
* WIDTH =       3                 *               *               *
*                                 *               *               *
* LEAST SQUARES SMOOTHING         *     0.0000499 *      0.0000387 *
* WIDTH =       3 DEGREE =   1  * *               *               *
**********************************************************************


SMOOTHING

      SMOOTHING FUNCTION--LEAST SQUARES
      NUMBER OF OBSERVATIONS        =        50


**********************************************************************
*                                 *   RESIDUAL    *   AVERAGE      *
*                                 *   STANDARD    *   ABSOLUTE     *
*                                 *   DEVIATION   *   RESIDUAL     *
*                                 * (DIVISOR=N-1) *  (DIVISOR=N)   *
**********************************************************************
* NO SMOOTHING                    *     0.0004291 *      0.0003570 *
* (RAW DATA)                      *               *               *
*                                 *               *               *
* MOVING AVERAGE SMOOTHING        *     0.0000499 *      0.0000387 *
* WIDTH =       3                 *               *               *
*                                 *               *               *
* LEAST SQUARES SMOOTHING         *     0.0000000 *      0.0000000 *
* WIDTH =       3 DEGREE =   3  * *               *               *
**********************************************************************
```

```
     SMOOTHING

          SMOOTHING FUNCTION--ROBUST (3RSR)
          NUMBER OF OBSERVATIONS        =        50
          DEFINING STRING               =

     *********************************************************************
     *                                *    RESIDUAL    *    AVERAGE      *
     *                                *    STANDARD    *    ABSOLUTE     *
     *                                *    DEVIATION   *    RESIDUAL     *
     *                                *  (DIVISOR=N-1) *   (DIVISOR=N)   *
     *********************************************************************
     * NO SMOOTHING                   *     0.0004291 *      0.0003570 *
     * (RAW DATA)                     *               *               *
     *                                *               *               *
     * MOVING AVERAGE SMOOTHING       *     0.0000499 *      0.0000387 *
     * WIDTH =      3                 *               *               *
     *                                *               *               *
     * ROBUST SMOOTHING (3RSR)        *     0.0000768 *      0.0000419 *
     *                                *               *               *
     *********************************************************************


     SMOOTHING

          SMOOTHING FUNCTION--MOVING MEDIAN
          NUMBER OF OBSERVATIONS        =        50


     *********************************************************************
     *                                *    RESIDUAL    *    AVERAGE      *
     *                                *    STANDARD    *    ABSOLUTE     *
     *                                *    DEVIATION   *    RESIDUAL     *
     *                                *  (DIVISOR=N-1) *   (DIVISOR=N)   *
     *********************************************************************
     * NO SMOOTHING                   *     0.0004291 *      0.0003570 *
     * (RAW DATA)                     *               *               *
     *                                *               *               *
     * MOVING AVERAGE SMOOTHING       *     0.0000499 *      0.0000387 *
     * WIDTH =      3                 *               *               *
     *                                *               *               *
     * MOVING MEDIAN SMOOTHING        *     0.0000495 *      0.0000240 *
     * WIDTH =      3                 *               *               *
     *********************************************************************
```