CRVFIT Version 1.0

29 October 2013

Doerry@aol.com

1. Introduction

`crvfit` is used to curve fit data to a curve of the form:

$$y = \sum_{i=0}^{n-1} C_i F_i(\acute{x})$$

$$\acute{x} = -1 + \frac{(x - x_{min})}{(x_{max} - x_{min})}$$

Where:

C_i = Curve coefficients of order *i*.

$F_i(\acute{x})$ = Characteristic Function of order *i*.

x_{min} = Minimum Value for x for evaluating the curve

x_{max} = Maximum Value for x for evaluating the curve

The sets of characteristic functions currently supported are:

`POLYNOMIAL:` $F_i(x) = x^i$

`LEGENDRE:` $F_i(x) = P_i(x)$

`CHEBYSHEV:` $F_i(x) = T_i(x)$

`LINEAR:` $F_0(x) = 1 \quad F_1(x) = x \quad n = 2$

2. Syntax

```
crvfit.exe xmlfilename [-oOutputFilename] [-d]
```

where:

`xmlfilename:` specifies the XML formatted filename with input data. Recommend using a `.xml` extension to allow viewing in a web browser.

OutputFilename: specifies file to hold the results. If not specified the results are shown on `stdout`. Recommend using a `.csv` extension on the filename to facilitate opening in Microsoft Excel

`-d: debug mode`

3. XML File Format

The XML file used as input has the following format:

```
<CURVEFIT>

<COMMENT> COMMENT elements are ignored </COMMENT>

<DATA> X Y </DATA>
<DATA> X Y </DATA>
<COMMENT> insert as many DATA elements as necessary </COMMENT>

<CURVE> LEGENDRE </CURVE>
<COMMENT> Other options are POLYNOMIAL, LINEAR and CHEBYSHEV
</COMMENT>

<NUMBER_TERMS> 5 </NUMBER_TERMS>
<COMMENT> Number of terms to generate for the series expansion.
If the CURVE element is LINEAR, should be set to 2 </COMMENT>

<XMIN> 0 </XMIN>
<COMMENT> Optional, mapped to -1 for series.  If not specified
then the minimum x value in the DATA elements is used </COMMENT>

<XMAX> 1 </XMAX>
<COMMENT> Optional, mapped to +1 for series.  If not specified
then the maximum x value in the DATA elements is used </COMMENT>

<CALCULATE>

<COMMENT> Optional, This section specifies for which values of x
the curve should be evaluated </COMMENT>

<X_VALUE> .5 </X_VALUE>

<COMMENT> Optional, Multiple X_VALUE elements can be specified.
The curve is evaluated for this value of x </COMMENT>

<X_PERCENTILE> .1 </X_PERCENTILE>

<COMMENT> Optional, Multiple X_PERCENTILE elements can be
specified.  The curve is evaluated at the value of x for the DATA
element with the largest x for which the fraction of other DATA
```

```
elements with a smaller value of x as compared to all the DATA
ELEMENTS is less than the specific X_PERCENTILE.  X_PERCENTILE
must be between 0 and 1.0 </COMMENT>

</CALCULATE>

</CURVEFIT>
```

4. Calculation Method

This program uses the pseudo-inverse method to calculate the characteristic function coefficients. In general, the following matrix equation holds:

$$Y = A(X)C$$

Where:

Y = vector of y values from the `DATA` elements (size m)

X = vector of $\acute{x}$ values calculated from the x values from the `DATA` elements and `XMIN` and `XMAX`. (size m)

m = number of `DATA` elements

n = number of terms in the series expansion

C = vector of coefficients C_i for the series expansion (size n)

$A(X)$ = matrix of characteristic functions evaluated at X (size m by n)

$$A(X) = \begin{vmatrix} F_0(\acute{x_0}) & F_1(\acute{x_0}) & F_2(\acute{x_0}) & \cdots & F_{n-1}(\acute{x_0}) \\ F_0(\acute{x_1}) & F_1(\acute{x_1}) & F_2(\acute{x_1}) & \cdots & F_{n-1}(\acute{x_1}) \\ F_0(\acute{x_2}) & F_1(\acute{x_2}) & F_2(\acute{x_2}) & \cdots & F_{n-1}(\acute{x_2}) \\ F_0(\acute{x_3}) & F_1(\acute{x_3}) & F_2(\acute{x_3}) & \cdots & F_{n-1}(\acute{x_3}) \\ \vdots & \vdots & \vdots & & \vdots \\ F_0(\acute{x_{m-1}}) & F_1(\acute{x_{m-1}}) & F_2(\acute{x_{m-1}}) & \cdots & F_{n-1}(\acute{x_{m-1}}) \end{vmatrix}$$

The following matrix algebra is used to generate the pseudo-inverse:

$$Y = AC$$
$$A^T Y = A^T AC$$
$$(A^T A)^{-1} A^T Y = C$$

The pseudo-inverse is the matrix $(A^T A)^{-1} A^T$. Where the T superscript indicates the transpose operator and the -1 superscript indicates matrix inversion. Hence C is calculated by multiplying the pseudo-inverse matrix by the Y vector. Once C is known, the curve can be evaluated for any value of x (mapped first to $\acute{x}$).

5. Example

Sample input:

```
<CURVEFIT>
<DATA>  0       2       </DATA>
<DATA>  0.1     2.339   </DATA>
<DATA>  0.2     2.752   </DATA>
<DATA>  0.3     3.233   </DATA>
<DATA>  0.4     3.776   </DATA>
<DATA>  0.5     4.375   </DATA>
<DATA>  0.6     5.024   </DATA>
<DATA>  0.7     5.717   </DATA>
<DATA>  0.8     6.448   </DATA>
<DATA>  0.9     7.211   </DATA>
<DATA>  1       8       </DATA>
<DATA>  1.1     8.809   </DATA>
<DATA>  1.2     9.632   </DATA>
<DATA>  1.3     10.463  </DATA>
<DATA>  1.4     11.296  </DATA>
<DATA>  1.5     12.125  </DATA>
<CURVE> LEGENDRE        </CURVE>
<NUMBER_TERMS> 6        </NUMBER_TERMS>
<CALCULATE>
<X_PERCENTILE> 0.1      </X_PERCENTILE>
<X_PERCENTILE> 0.2      </X_PERCENTILE>
<X_PERCENTILE> 0.3      </X_PERCENTILE>
<X_PERCENTILE> 0.4      </X_PERCENTILE>
<X_PERCENTILE> 0.5      </X_PERCENTILE>
<X_PERCENTILE> 0.6      </X_PERCENTILE>
<X_PERCENTILE> 0.7      </X_PERCENTILE>
<X_PERCENTILE> 0.8      </X_PERCENTILE>
<X_PERCENTILE> 0.9      </X_PERCENTILE>
<X_PERCENTILE> 1        </X_PERCENTILE>
<X_VALUE>       0       </X_VALUE>
<X_VALUE>       0.5     </X_VALUE>
<X_VALUE>       1       </X_VALUE>
<X_VALUE>       1.5     </X_VALUE>
</CALCULATE>
</CURVEFIT>
```

Sample output:

```
CRVFIT ,1.0

 XML File , test1.xml
 xmin , 0.000000
 xmax , 1.500000
 POLYNOMIAL <> Number Terms , 4
  Coefficient 0 , 6.504185
  Coefficient 1 , 4.982395
  Coefficient 2 , 1.196432
  Coefficient 3 , 0.270011


 DATA

 x_percentile ,     x    ,      y
   0.000000 ,   0.000000 ,   2.275061
   0.062500 ,   0.100000 ,   3.337803
   0.125000 ,   0.200000 ,   3.203263
```

```
   0.187500 ,    0.300000 ,    3.654958
   0.250000 ,    0.400000 ,    4.393186
   0.312500 ,    0.500000 ,    5.159604
   0.375000 ,    0.600000 ,    5.841717
   0.437500 ,    0.700000 ,    5.761142
   0.500000 ,    0.800000 ,    7.184431
   0.562500 ,    0.900000 ,    7.290799
   0.625000 ,    1.000000 ,    8.026713
   0.687500 ,    1.100000 ,    9.552981
   0.750000 ,    1.200000 ,    9.710478
   0.812500 ,    1.300000 ,   11.086442
   0.875000 ,    1.400000 ,   11.963528
   0.937500 ,    1.500000 ,   12.856624
```

```
CALCULATE X
```

```
         x     ,       y
   0.000000 ,    2.448212
   0.100000 ,    2.908996
   0.200000 ,    3.387359
   0.300000 ,    3.887142
   0.400000 ,    4.412183
   0.500000 ,    4.966323
   0.600000 ,    5.553403
   0.700000 ,    6.177263
   0.800000 ,    6.841742
   0.900000 ,    7.550681
   1.000000 ,    8.307921
   1.100000 ,    9.117300
   1.200000 ,    9.982660
   1.300000 ,   10.907840
   1.400000 ,   11.896681
   1.500000 ,   12.953023
```

```
CALCULATE X PERCENTILE
```

```
 x_percentile ,      x     ,       y
   0.250000 ,    0.400000 ,    4.412183
   0.500000 ,    0.800000 ,    6.841742
   0.750000 ,    1.200000 ,    9.982660
```

Comparison Graph:

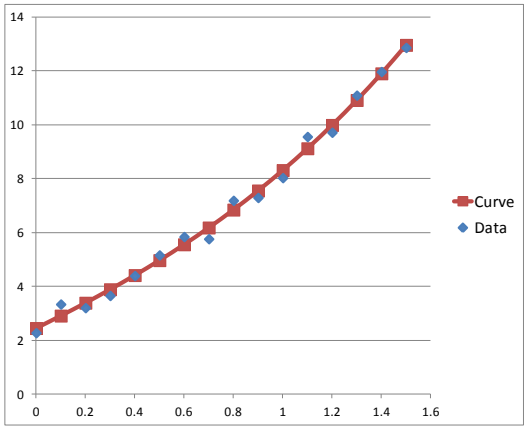

Figure 1: Comparison of Data to Curve

6. Recommendations

One way to create the XML file is to use Microsoft Excel as a means to structure the document, copy and paste the contents into a NOTEPAD text file, then replace the `.txt` file extension with `.xml`.

The `.xml` document can be viewed with a web browser to ensure the XML is well-formed.

If the output file is provided with a `.csv` file extension, then the file can be directly opened in Microsoft Excel for further processing.

7. Notice