StreamBase Polynomial Regression Examples

Import non-linear regression library

In[7]:=
```
Needs["NonlinearRegression`"]
```

Example 1

Massage the data into the proper format where the data is from random points on a sinusoid.

In[8]:=
```
x = {2, 6, 10, 14, 18};
y = {-0.99, -0.7, 0.58, 0.41, -0.66};
data = Table[{x[[i]], y[[i]]}, {i, 1, Length[x]}];
```

Regression with highest monomial degree 2

In[11]:=
```
NonlinearRegress[data, a + b * z + c * z^2, {a, b, c}, z]
```

Out[11]=
```
{BestFitParameters → {a → -1.98039, b → 0.416571, c → -0.0186161},

 ParameterCITable →
   |  Estimate     Asymptotic SE  CI
 a | -1.98039     0.716137       {-5.06168, 1.1009}
 b | 0.416571     0.169044       {-0.310768, 1.14391}
 c | -0.0186161   0.00822505     {-0.0540056, 0.0167735} ,

 EstimatedVariance → 0.242463,

 ANOVATable →
                          DF     SumOfSq      MeanSq
   Model                  3      1.92527      0.641758
   Error                  2      0.484926     0.242463,
   Uncorrected Total      5      2.4102
   Corrected Total        4      2.04028

 AsymptoticCorrelationMatrix →
   ( 1.         -0.885188  0.781001
     -0.885188  1.         -0.973124
     0.781001   -0.973124  1.        ),

 FitCurvatureTable →
                                 | Curvature
   Max Intrinsic                 | 0
   Max Parameter-Effects         | 0
   95. % Confidence Region       | 0.22843 }
```

Example 2

Massage the data into the proper format where the data is US Census data relating wage to median income.

```
In[12]:= Ages = {27, 32, 37, 42, 47, 52, 57, 62, 67, 72, 77};
         MedianIncome =
           {47358, 55077, 61782, 62578, 64802, 66244, 61174, 52428, 40296, 31654, 23230};
         censusData = Table[{Ages[[i]], MedianIncome[[i]]}, {i, 1, Length[Ages]}];
```

Regression with highest monomial degree 4

```
In[15]:= NonlinearRegress[censusData, a + b * z + c * z^2 + d * z^3 + f * z^4, {a, b, c, d, f}, z]
```

Out[15]=

$\{$BestFitParameters $\to \{a \to 136498.,\ b \to -11453.8,\ c \to 471.088,\ d \to -7.20033,\ f \to 0.0359142\}$,

ParameterCITable $\to$

	Estimate	Asymptotic SE	CI
a	136498.	80372.2	{−60166., 333162.}
b	−11453.8	6933.63	{−28419.7, 5512.2}
c	471.088	214.896	{−54.7449, 996.92}
d	−7.20033	2.84718	{−14.1671, −0.233543}
f	0.0359142	0.0136621	{0.00248436, 0.0693441}

,

EstimatedVariance $\to 3.00276 \times 10^6$,

ANOVATable $\to$

	DF	SumOfSq	MeanSq
Model	5	3.12352×10^{10}	6.24703×10^9
Error	6	1.80166×10^7	3.00276×10^6
Uncorrected Total	11	3.12532×10^{10}	
Corrected Total	10	2.06575×10^9	

,

AsymptoticCorrelationMatrix $\to$

$$\begin{pmatrix} 1. & -0.997054 & 0.989053 & -0.977358 & 0.963239 \\ -0.997054 & 1. & -0.997378 & 0.990342 & -0.980143 \\ 0.989053 & -0.997378 & 1. & -0.997735 & 0.991708 \\ -0.977358 & 0.990342 & -0.997735 & 1. & -0.998079 \\ 0.963239 & -0.980143 & 0.991708 & -0.998079 & 1. \end{pmatrix}$$

,

FitCurvatureTable $\to$

	Curvature
Max Intrinsic	0
Max Parameter-Effects	0
95. % Confidence Region	0.477417

$\}$