41. 曲线拟合的瑞士军刀 — 1stOpt 神经网络拟合特例演示

41.1 引言

瑞士军刀以小巧、精致、功能丰富且强悍而闻名于世逾百年，常被隐喻为解决疑难问题的利器和工具。非线性曲线拟合是日常科学研究中最常遇到的数学问题之一，尤其是在拟合模型公式未知而只有数据，同时这些数据所展示的图形关系又表现的异于常规，对普通人而言很难根据常规知识猜测确定出可能的模型公式。虽然 1stOpt 有一个快速拟合工具箱，可以根据数据自动搜寻匹配出最佳公式，但对异于常规的数据所搜寻得出的模型公式，往往其拟合计算结果仍然很差，难以满足高质量拟合要求。遇到这种情况，该如何处理？有无其它好的解决方法？此等“万般无奈”的困境下，1stOpt 的神经网络拟合功能或工具箱无疑就是最佳选择了，完全满足①小巧：一句代码实现；②强大：几乎可以拟合任意形式的数据曲线，同时还能给出具体的模型公式表达式，这不正是“众里寻他千百度，蓦然回首，那人却在，灯火阑珊处”的曲线拟合“瑞士军刀”吗！

以六个实际数据拟合问题为例，演示 1stOpt 神经网络拟合功能及工具箱在处理特殊形式的曲线拟合所展示的强大与便捷。当然核心基础还要同时归功于 1stOpt 自身独一无二的全局优化算法，否则也只能是“巧妇难为无米之炊”，望洋兴叹了。

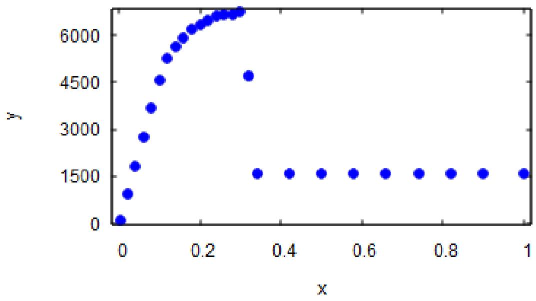
41.2 案例数据

六个案例实际数据见表 41-1，均只有一个自变量 x 和一个因变量 y，数据图形都比较特殊，参见图 41-1，具体模型公式均未知。

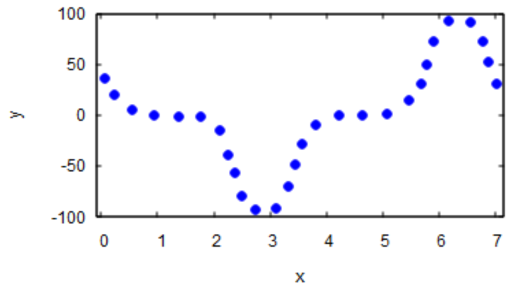
表 41-1：案例数据

案例1	x	0.002,0.02,0.04,0.06,0.08,0.1,0.12,0.14,0.16,0.18,0.2,0.22,0.24,0.26,0.28,0.3,0.32,0.34,0.42,0.5,0.58,0.66,0.74,0.82,0.9,1
	y	2.0295,920.295,1840.59,2760.89,3676.04,4549.25,5255.34,5624.14,5922.58,6173.37,6348.77,6461.03,6612.06,6673.27,6663.94,6724.78,4702.07,1581.33,1581.33,1581.33,1581.33,1581.33,1581.33,1581.33,1581.33,1581.33
案例2	x	0.058,0.233,0.548,0.946,1.378,1.775,2.102,2.254,2.370,2.499,2.744,3.118,3.328,3.445,3.573,3.807,4.216,4.648,5.080,5.466,5.676,5.793,5.909,6.174,6.552,6.774,6.891,7.019
	y	36.67,20.36,5.31,0.29,-0.97,-2.22,-15.39,-39.23,-56.17,-80.01,-93.81,-91.30,-69.97,-49.27,-27.94,-9.12,-0.34,0.29,1.54,14.09,30.40,49.22,72.43,92.65,91.25,71.80,52.35,30.40
案例3	x	0.025,0.153,0.294,0.597,1.123,1.613,2.081,2.606,3.062,3.295,3.400,3.482,3.576,3.704,3.996,4.545,5.129,5.713,6.262,6.647,6.811,6.904,6.986
	y	0.017,-0.411,-0.616,-0.870,-0.914,-0.933,-0.945,-0.908,-0.808,-0.523,-0.331,-0.020,0.370,0.557, 0.867,0.947,0.966,0.954,0.916,0.687,0.358,0.085,-0.138
案例4	x	1.925,2.053,2.194,2.497,3.062,3.295,3.400,3.482,3.622,3.821,4.253,4.837,5.421,6.005,6.495,6.741,6.857,6.951,7.079,7.194,7.497,8.062,8.295;
	y	0.017,-0.411,-0.616,-0.870,-0.808,-0.523,-0.331,-0.020,0.395,0.761,0.916,0.960,0.960, 0.941, 0.823,0.532,0.203,0.011,-0.405,-0.616,-0.870,-0.808,-0.523;
案例5	x	1.866,2.007,2.112,2.31,2.742,3.202,3.354,3.447,3.517,3.704,3.996,4.253,4.545,4.837,5.129,5.421,5.713, 6.005,6.262,6.495,6.647,6.811,6.904,6.986,7.112,7.31,7.497,7.742,8.062,8.202,8.295;
	y	0.203,-0.244,-0.511,-0.746,-0.914,-0.672,-0.418,-0.138,0.073,0.557,0.867,0.916,1.2,2.0,2.5,2.0, 1.2,0.941,0.916,0.823,0.687,0.358,0.085,-0.138,-0.511,-0.746,-0.870,-0.914,-0.808,-0.672,-0.523;
案	x	-5,-4.5,-4,-3.5,-3,-2.4,-1.6,-1.5,-1.4,-1.3,-1.2,-1.1,-0.9,-0.7,-0.4,-0.3,-0.2,-0.1,0,0.1,0.2,0.3,0.4, 0.7, 0.9,1.1,1.2,1.3,1.4,1.5,1.6,2.4,2.9,3.4,3.9,4.4,5;

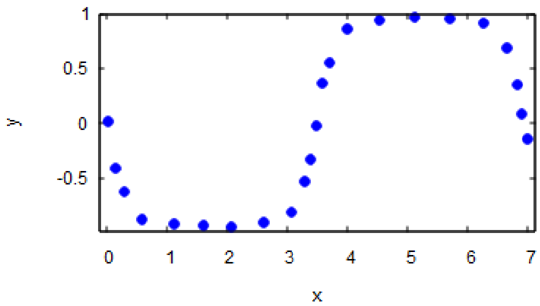
例 6	y	0,0,0,0,0,0,0.002,0.082,0.427,0.783,0.946,0.991,1.000,1.000,0.999,0.988,0.933,0.800,0.5,0.800,0.933,0.988,0.999,1.000,1.000,0.991,0.946,0.783,0.427,0.082,0.002,0,0,0,0,0,0;



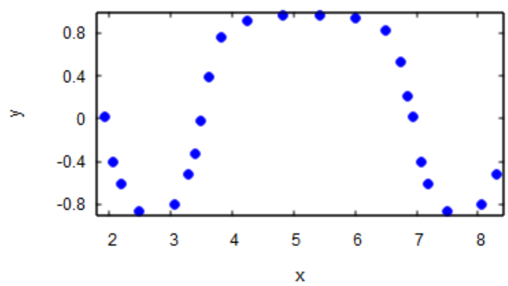
案例-1



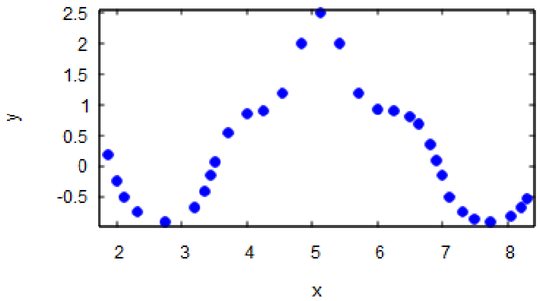
案例-2



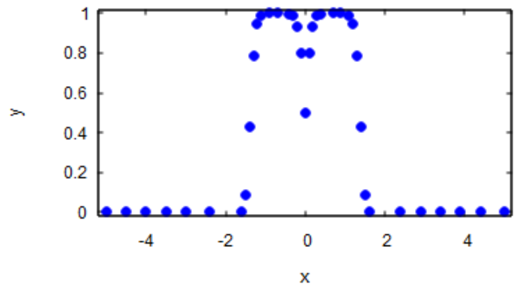
案例-3



案例-4



案例-5



案例-6

图 41-1 案例数据散点曲线图

41.3 神经网络结构

1stOpt 神经网络拟合代码可以通过两种方式产生，一种是可视化的神经网络拟合工具箱，如图 41-2 示，可以可视化设定神经网络的所有要素如隐含层数、隐含层神经元数、激活函数类型、连接方式、数据是否均一化等，最终输出拟合代码；第二种方式是已经熟悉了神经网络拟合代码所代表的含义，直接在代码里通过快捷键输入拟合代码，如图 41-3 示，有两个快捷命令“NNFit”及“NNFit1”，对应的拟合缺省命令分别为：

NNFit： NNFit(NS=[1-3-1], TF=[1,1], DN=[1,1], NC=1, CT=1, Code=1, DC=[]);

NNFit1： NNFit(NS=[1-3-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]);

说明：

- ✧ NS=[1-3-1]：输入与输出层均只有一个神经元，一个隐含层含三个神经元；
- ✧ TF=[1,1]：隐含层和输出层激活函数均为 Sigmoid；
- ✧ TF=[1,4]：隐含层激活函数为 Sigmoid，输出层激活函数为线性 Linear 函数；

- ✧ DN=[1,1]：输入输出层数据全部均一化处理；
- ✧ DN=[0,0]：输入输出层数据全部不做均一化处理，保持原数据；

上述缺省拟合命令可以根据实际情况修改，大多数情况下一个隐含层足够，至于隐含层应该有多少个神经元，视具体问题而定，一般在满足拟合精度的前提下神经元数越少越好，太多会造成过拟合现象，反之，太少又会造成欠拟合，难以满足精度要求。本文所涉及的六个案例，其各自的神经网络结构见图 41-4，对应的拟合公式及快捷拟合命令见表 41-2.

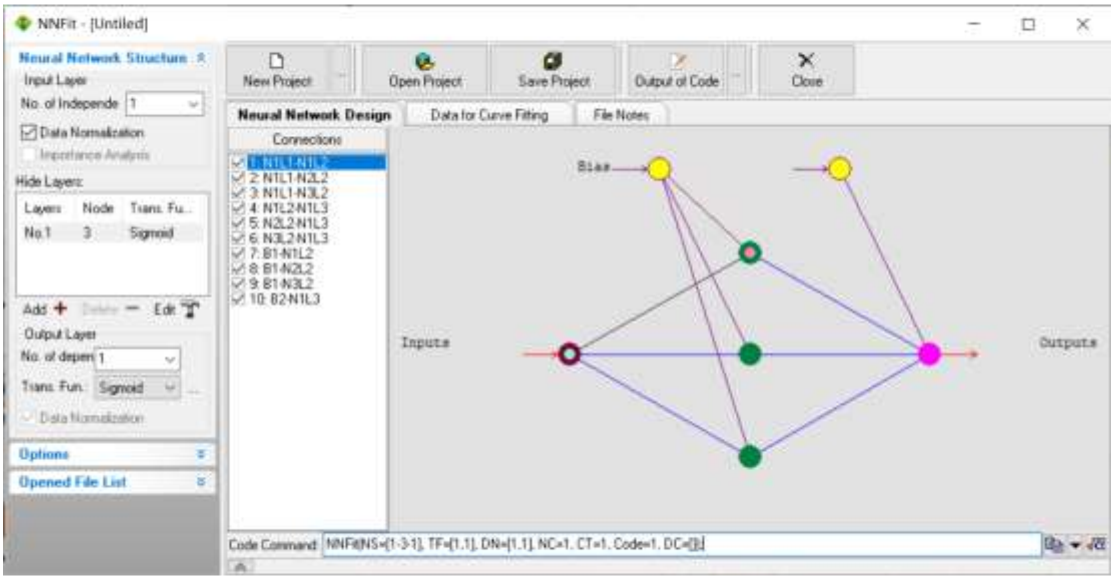


图 41-2 神经网络拟合工具箱

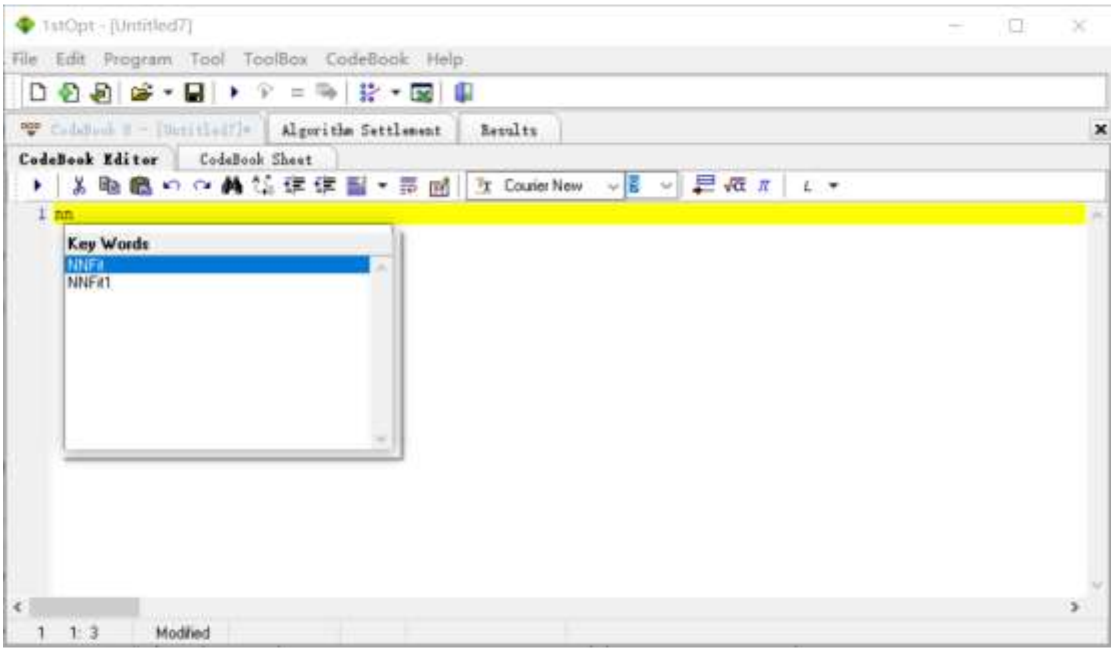
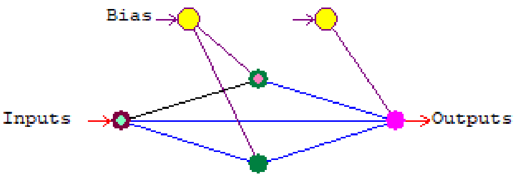
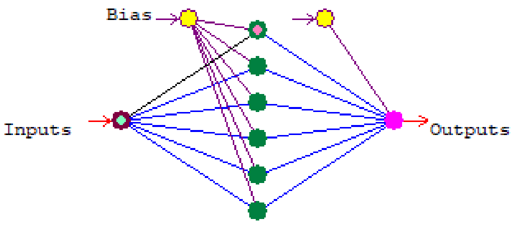


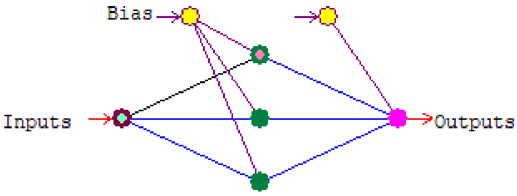
图 41-3 神经网络拟合命令快捷输入



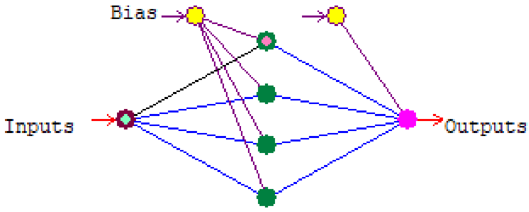
案例-1



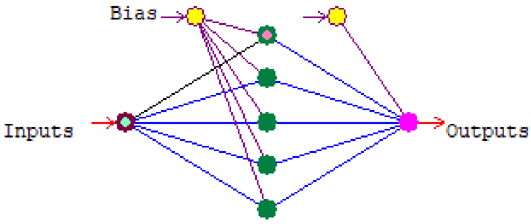
案例-2



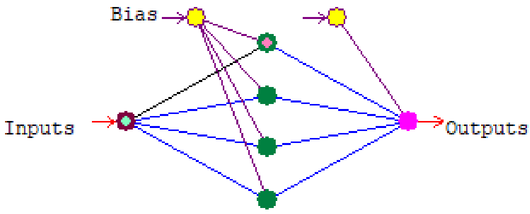
案例-3



案例-4



案例-5



案例-6

图 41-4 各案例神经网络结构

表 41-2 各案例神经网络拟合快捷命令及对应的模型公式

案例1	NNFit(NS=[1-2-1], TF=[1,4], DN=[0,0], NC=2, CT=1, Code=1, DC=[]);
	$y=\frac{p_3}{1+\exp\left(-(p_1\cdot x+p_6)\right)}+\frac{p_4}{1+\exp\left(-(p_2\cdot x+p_7)\right)}+p_8+p_5\cdot x$
案例2	NNFit(NS=[1-6-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]);
	$y=\frac{p_7}{1+\exp\left(-(p_1\cdot x+p_{13})\right)}+\frac{p_8}{1+\exp\left(-(p_2\cdot x+p_{14})\right)}+\frac{p_9}{1+\exp\left(-(p_3\cdot x+p_{15})\right)}+\frac{p_{10}}{1+\exp\left(-(p_4\cdot x+p_{16})\right)}+\frac{p_{11}}{1+\exp\left(-(p_5\cdot x+p_{17})\right)}+\frac{p_{12}}{1+\exp\left(-(p_6\cdot x+p_{18})\right)}+p_{19}$
案例3	NNFit(NS=[1-3-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]);
	$y=\frac{p_4}{1+\exp\left(-(p_1\cdot x+p_7)\right)}+\frac{p_5}{1+\exp\left(-(p_2\cdot x+p_8)\right)}+\frac{p_6}{1+\exp\left(-(p_3\cdot x+p_9)\right)}+p_{10}$
案	NNFit(NS=[1-4-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]);

例4	$$y=\frac{p_5}{1+\exp\big(-(p_1\cdot x+p_9)\big)}+\frac{p_6}{1+\exp\big(-(p_2\cdot x+p_{10})\big)}+\frac{p_7}{1+\exp\big(-(p_3\cdot x+p_{11})\big)}+\frac{p_8}{1+\exp\big(-(p_4\cdot x+p_{12})\big)}+p_{13}$$
案例5	NNFit(NS=[1-5-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]); $$y=\frac{p_6}{1+\exp\big(-(p_1\cdot x+p_{11})\big)}+\frac{p_7}{1+\exp\big(-(p_2\cdot x+p_{12})\big)}+\frac{p_8}{1+\exp\big(-(p_3\cdot x+p_{13})\big)}+\frac{p_9}{1+\exp\big(-(p_4\cdot x+p_{14})\big)}+\frac{p_{10}}{1+\exp\big(-(p_5\cdot x+p_{15})\big)}+p_{16}$$
案例6	NNFit(NS=[1-4-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]); $$y=\frac{p_5}{1+\exp\big(-(p_1\cdot x+p_9)\big)}+\frac{p_6}{1+\exp\big(-(p_2\cdot x+p_{10})\big)}+\frac{p_7}{1+\exp\big(-(p_3\cdot x+p_{11})\big)}+\frac{p_8}{1+\exp\big(-(p_4\cdot x+p_{12})\big)}+p_{13}$$

41.4 计算代码及结果

各案例计算代码、计算结果及最终拟合结果对比图如下，可以看出拟合结果都是非常好的。

案例-1代码

```
Function NNFit(NS=[1-2-1], TF=[1,4], DN=[0,0], NC=2, CT=1, Code=1, DC=[]);
Data;
x=0.002,0.02,0.04,0.06,0.08,0.1,0.12,0.14,0.16,0.18,0.2,0.22,0.24,0.26,0.28,0.3,0.32,0.34,0.42,0.5,0.58,0.66,0.74,0.82,0.9,1;
y=2.0295,920.295,1840.59,2760.89,3676.04,4549.25,5255.34,5624.14,5922.58,6173.37,6348.77,6461.03,6612.06,6673.27,6663.94,6724.78,4702.07,1581.33,1581.33,1581.33,1581.33,1581.33,1581.33,1581.33,1581.33,1581.33;
```

案例-1结果

```
Sum Squared Error (SSE): 37100.3310237717
Root of Mean Square Error (RMSE): 37.7748038857464
Correlation Coef. (R): 0.999862475582419
R-Square: 0.999724970077804
Adjusted R-Square: 0.999701054432396
Determination Coef. (DC): 0.999724970077804
Chi-Square: 35.6295512917898
F-Statistic: 9347.0606071207

Parameter	Best Estimate
---------	-------------
p1	-20.5063832358883
p2	-392.695394569989
p3	-9735.23368413293
p4	5156.64199091371
p5	-10.3506554410191
p6	0.838586420308152
p7	126.110546308293
p8	1590.73873715065
```

案例-2代码

```
Function NNFit(NS=[1-6-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]);
Data;
x=0.058,0.233,0.548,0.946,1.378,1.775,2.102,2.254,2.370,2.499,2.744,3.118,3.328,3.445,3.573,3.807,4.216,4.6
```

```
48,5.080,5.466,5.676,5.793,5.909,6.174,6.552,6.774,6.891,7.019;
y=36.67,20.36,5.31,0.29,-0.97,-2.22,-15.39,-39.23,-56.17,-80.01,-93.81,-91.30,-69.97,-49.27,-27.94,-9.12,-
0.34,0.29,1.54,14.09,30.40,49.22,72.43,92.65,91.25,71.80,52.35,30.40;
```

案例-2结果

```
Sum Squared Error (SSE): 47.7419990274329
Root of Mean Square Error (RMSE): 1.30578382354695
Correlation Coef. (R): 0.999672650036559
R-Square: 0.999345407231116
Adjusted R-Square: 0.999293039809606
Determination Coef. (DC): 0.999345407231117
Chi-Square: -0.136439556890028
F-Statistic: 763.333673872321

Parameter  Best Estimate
---------  -------------
p1         -7.45925331333407
p2         -6.02760393193255
p3         11344.3661260844
p4         -6.99380393513509
p5         4.54688471731932
p6         -6.61636164622797
p7         99.3393436818711
p8         118.688806591866
p9         159.438734291684
p10        -100.473361477677
p11        -102.432826019644
p12        -101.762517036535
p13        17.3520053648941
p14        41.8787068631259
p15        1836.49872185009
p16        24.09825119692
p17        0.290609192585508
p18        38.3353041349946
p19        -73.5066791694304
```

案例-3代码

```
Function NNFit(NS=[1-3-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]);
Data;
x=0.025,0.153,0.294,0.597,1.123,1.613,2.081,2.606,3.062,3.295,3.400,3.482,3.576,3.704,3.996,4.545,5.129,5.7
13,6.262,6.647,6.811,6.904,6.986;
y=0.017,-0.411,-0.616,-0.870,-0.914,-0.933,-0.945,-0.908,-0.808,-0.523,-0.331,-
0.020,0.370,0.557,0.867,0.947,0.966,0.954,0.916,0.687,0.358,0.085,-0.138;
```

案例-3结果

```
Sum Squared Error (SSE): 0.0172811972662589
Root of Mean Square Error (RMSE): 0.0274108811036935
Correlation Coef. (R): 0.999224673789575
R-Square: 0.998449948709882
Adjusted R-Square: 0.99829494358087
Determination Coef. (DC): 0.998449948709882
Chi-Square: 0.00405200904568994
F-Statistic: 930.424340235396

Parameter  Best Estimate
---------  -------------
p1         -6.95867626794223
p2         -6.65517532715326
p3         -4.43948102844079
p4         -1.87102497778266
p5         1.74210592740286
p6         431058.022943748
p7         24.277029347241
p8         45.9890810488087
```

```
p9          -12.9257201691792
p10         -0.796863447193001
```

案例-4代码

```
Function NNFit(NS=[1-4-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]);
Data;
x=1.925,2.053,2.194,2.497,3.062,3.295,3.400,3.482,3.622,3.821,4.253,4.837,5.421,6.005,6.495,6.741,6.857,6.9
51,7.079,7.194,7.497,8.062,8.295;
y=0.017,-0.411,-0.616,-0.870,-0.808,-0.523,-0.331,-
0.020,0.395,0.761,0.916,0.960,0.960,0.941,0.823,0.532,0.203,0.011,-0.405,-0.616,-0.870,-0.808,-0.523;
```

案例-4结果

```
Sum Squared Error (SSE): 0.00912808990623831
Root of Mean Square Error (RMSE): 0.0199216835178223
Correlation Coef. (R): 0.999537467873735
R-Square: 0.999075149683437
Adjusted R-Square: 0.998982664651781
Determination Coef. (DC): 0.999075149683435
Chi-Square: -0.0953389564005932
F-Statistic: 900.213309292852

Parameter   Best Estimate
---------   -------------
p1          -4.6465596953849
p2          -6.86949975298208
p3          -6.19639294895301
p4          -4.98422914957766
p5          36.8432635190473
p6          -1.86032122430374
p7          1.88671902567172
p8          -741.362446220397
p9          5.28801786482015
p10         24.03915355189
p11         42.9930531459857
p12         48.8293107094801
p13         740.42296077957
```

案例-5代码

```
Function NNFit(NS=[1-5-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]);
Data;
x=1.866,2.007,2.112,2.31,2.742,3.202,3.354,3.447,3.517,3.704,3.996,4.253,4.545,4.837,5.129,5.421,5.713,6.00
5,6.262,6.495,6.647,6.811,6.904,6.986,7.112,7.31,7.497,7.742,8.062,8.202,8.295;
y=0.203,-0.244,-0.511,-0.746,-0.914,-0.672,-0.418,-
0.138,0.073,0.557,0.867,0.916,1.2,2.0,2.5,2.0,1.2,0.941,0.916,0.823,0.687,0.358,0.085,-0.138,-0.511,-0.746,-
0.870,-0.914,-0.808,-0.672,-0.523;
```

案例-5结果

```
Sum Squared Error (SSE): 0.0695566158134452
Root of Mean Square Error (RMSE): 0.0473683628641915
Correlation Coef. (R): 0.998705726415323
R-Square: 0.997413127974758
Adjusted R-Square: 0.997228351401526
Determination Coef. (DC): 0.997413127861668
Chi-Square: -0.000163900707998824
F-Statistic: 385.570959071078

Parameter   Best Estimate
---------   -------------
p1          1.4881630655839
p2          -1.13093125088893
p3          -2.4590246750242
```

p4	-2.61730593350181
p5	-4.7969937936172
p6	64.7929054077903
p7	2652.66628915952
p8	41.6075374273339
p9	14.5019582002414
p10	5.44973543758652
p11	-8.19345304915228
p12	-4.67640725269131
p13	13.4618325715375
p14	17.7009661157683
p15	20.2751454982163
p16	-64.6704442249966

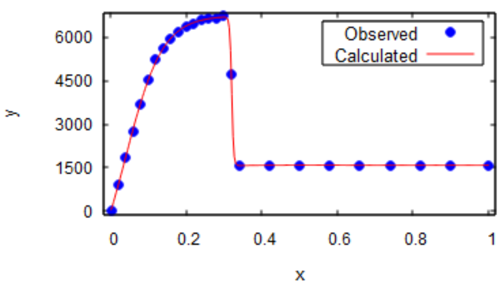
案例-6代码

```
Function NNFit(NS=[1-4-1], TF=[1,4], DN=[0,0], NC=1, CT=1, Code=1, DC=[]);
Data;
x=-5,-4.5,-4,-3.5,-3,-2.4,-1.6,-1.5,-1.4,-1.3,-1.2,-1.1,-0.9,-0.7,-0.4,-0.3,-0.2,-
0.1,0,0.1,0.2,0.3,0.4,0.7,0.9,1.1,1.2,1.3,1.4,1.5,1.6,2.4,2.9,3.4,3.9,4.4,5;
y=0,0,0,0,0,0,0.002,0.082,0.427,0.783,0.946,0.991,1.000,1.000,0.999,0.988,0.933,0.800,0.5,0.800,0.933,0.988,
0.999,1.000,1.000,0.991,0.946,0.783,0.427,0.082,0.002,0,0,0,0,0,0;
```

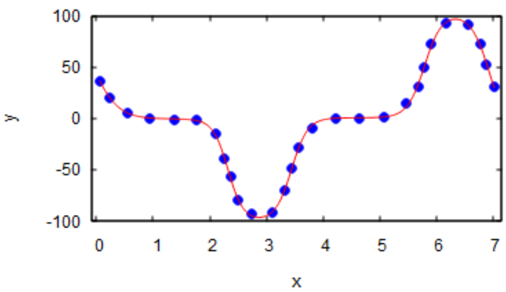
案例-6结果

Sum Squared Error (SSE): 0.00471743736955039
Root of Mean Square Error (RMSE): 0.0112915148357138
Correlation Coef. (R): 0.999682289576457
R-Square: 0.999364680092826
Adjusted R-Square: 0.999327308333581
Determination Coef. (DC): 0.999364680092826
Chi-Square: -0.0192723131953977
F-Statistic: 3146.02030978668

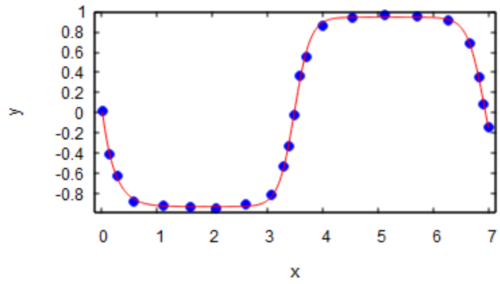
Parameter	Best Estimate
---------	-------------
p1	20.1136617022308
p2	20.113660246729
p3	17.6455646159104
p4	17.645566246348
p5	75.6148459507437
p6	-75.6148459950451
p7	0.996214534699603
p8	-0.996214489455113
p9	-0.0128863670640285
p10	0.0128710700132883
p11	24.3523927052369
p12	-24.352395093218
p13	-0.00403115010613218



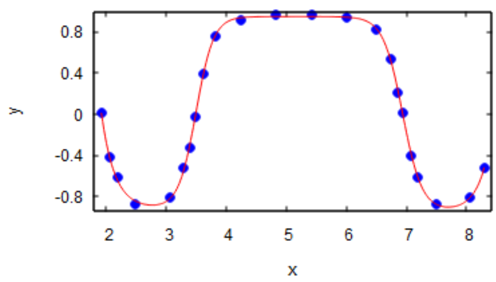
案例-1



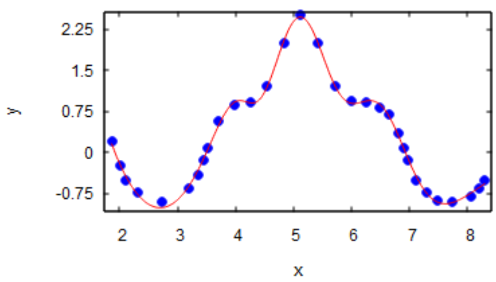
案例-2



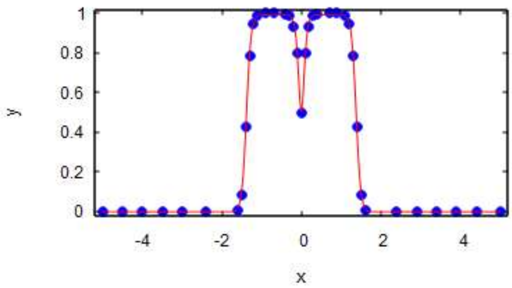
案例-3



案例-4



案例-5



案例-6

图 41-5 各案例拟合计算结果对比图

41.5 小结

在拟合公式未知及数据形状特殊的情况下如何获得理想的模型公式？基于实际案例，1stOpt 的神经网络拟合功能及工具箱无疑给出了很好的答案，几近完美的拟合结果展示了曲线拟合瑞士军刀的强大威力。当然，十全十美的东西世界上从来就不存在，神经网络虽然强悍，但也存在一些不足和短板，比如对简单的多周期三角函数数据拟合结果就差强人意，这也实属正常，如同瑞士军刀再威力无比无法替代枪炮一样。总而言之，在解决曲线拟合陷入困境时，1stOpt 的神经网络拟合工具箱是一个非常值得考虑的工具选项，除了简单且强大，还能实时给出对应的具体模型公式表达式，方便迁移和扩展，有着极大的应用潜力。