36. Matlab 拟合工具箱内置拟合函数功效测试－与 1stOpt 对比

36.1 引言

之前文章介绍了两款知名数据分析软件Origin和SigmaPlot内置拟合函数存在的问题，而作为公认的科学数值计算“盟主”的Matlab，其曲线拟合工具箱同样内置了一些函数，其拟合效果如何？是否同样存在“不需猜初值但却效果不佳”的问题？在此进行简单的实战验证，并同时与1stOpt结果进行对比。

36.2 Matlab 拟合工具箱内置拟合函数

如下图示，Matlab拟合工具箱内置函数除了自定义（Custom Equation）、插值（Interpolant）和平滑样条（Smoothing spline）外，主要还包括：

1) Exponential：指数函数；
2) Fourier：傅里叶函数；
3) Gausian：高斯函数；
4) Polynomial：多项式函数；
5) Power：幂函数；
6) Rational：有理函数；
7) Sum of Sin：正弦和函数；
8) Weibull：温布尔函数。

上述8个内置函数中，Polynomial和Power函数相对简单，仅选取其余6个进行计算对比和验证。

对于内置拟合函数的初值确定，按MathWork官方的说法“Starting points for some library models are determined heuristically”，即模型初值是通过启发式方法确定的，实际上与Origin和SigmaPlot确定初值的方式类似：软件本身提前“知道”每个内置函数中各个参数的实际几何物理意义，因此根据用户的实际数据在拟合计算之前就可以在后台估算出大概合理的参数初值，进而在此基础上迭代计算得到最终正确结果。

选用内置拟合函数是否就可真正免除用户猜初值的麻烦而获得正确结果？实例验证如下。

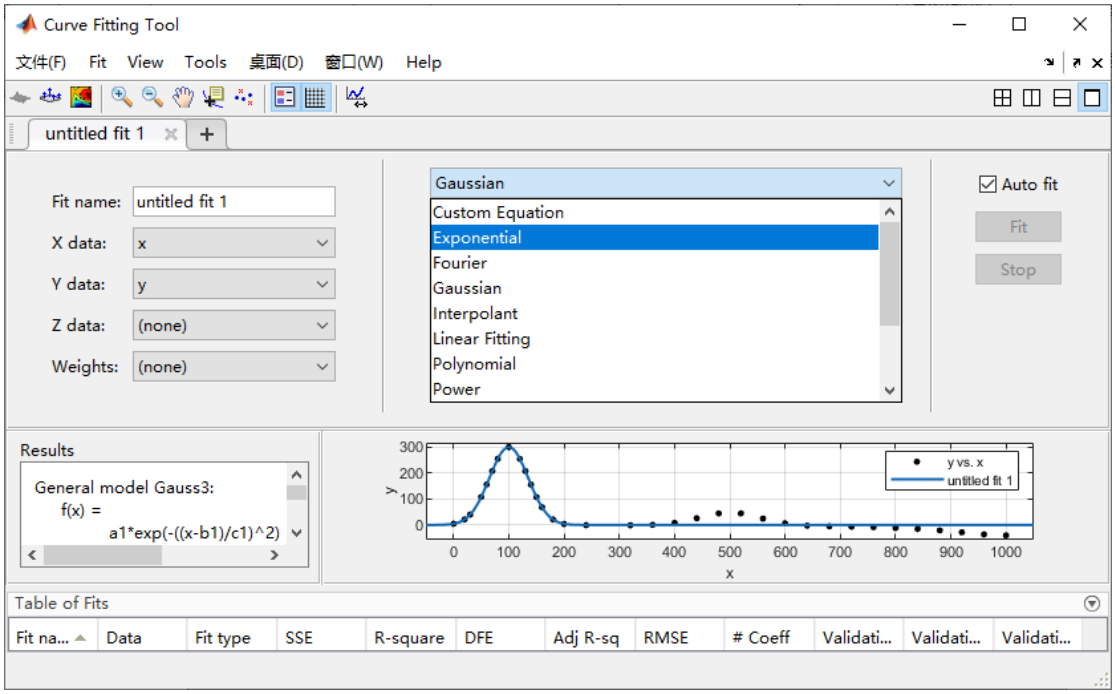

图 36.1 Matlab 拟合工具箱

Fit Category	Curves	Surfaces
Regression Models		
Polynomial	Yes (up to degree 9)	Yes (up to degree 5)
Exponential	Yes	
Fourier	Yes	
Gaussian	Yes	
Power	Yes	
Rational	Yes	
Sum of Sine	Yes	
Weibull	Yes	

图 36.2 Matlab 拟合工具箱内置拟合函数

36.3 Matlab 内置拟合函数实例验证

36.3.1 Exponential：指数函数

数据

x	0,1,1.5,1.6,1.8,2,2.25,2.3,2.4,2.55,2.61,2.7,2.75,2.8,2.9,2.94,3
y	0.17,0.2,0.99,1.56,3.55,4.52,11.14,13.27,16,22.65,25.3,29.57,31.65,33.91,39.13,40.6,42.55

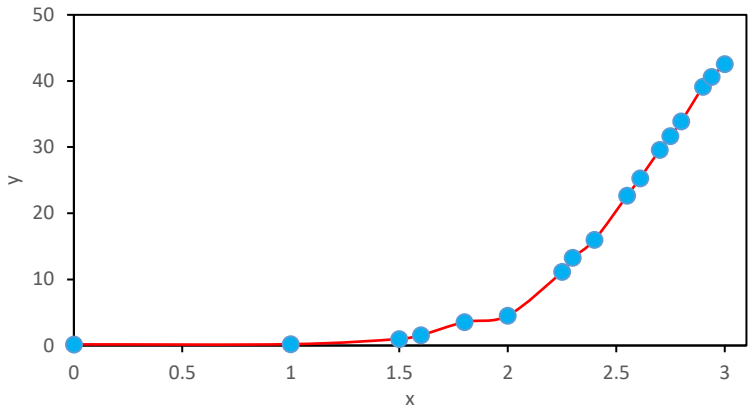

图 36.3 指数函数（Exponential）验证数据图

函数形式

Exp1	$y = a \cdot \exp(b \cdot x)$
Exp2	$y = a \cdot \exp(b \cdot x) + c \cdot \exp(d \cdot x)$

Matlab即可在可视化拟合工具箱进行，也可以用如下代码：

Matlab代码：

```
clear;
clc;
close all;

% Curve Fit
x=[0,1,1.5,1.6,1.8,2,2.25,2.3,2.4,2.55,2.61,2.7,2.75,2.8,2.9,2.94,3];
y=[0.17,0.2,0.99,1.56,3.55,4.52,11.14,13.27,16,22.65,25.3,29.57,31.65,33.91,39.13,40.6,42.55];

theFit=fit(x',y','Exp2')
coeffvalues(theFit)
plot(theFit,x,y)
```

1stOpt代码：

```
Variable x,y;
Function y=a*exp(b*x);                    //Exp1
         //y=a*exp(b*x) + c*exp(d*x); //Exp2
Data;
z=[0,1,1.5,1.6,1.8,2,2.25,2.3,2.4,2.55,2.61,2.7,2.75,2.8,2.9,2.94,3];
y=[0.17,0.2,0.99,1.56,3.55,4.52,11.14,13.27,16,22.65,25.3,29.57,31.65,33.91,39.13,40.6,42.55];
```

Matlab与1stOpt计算结果

函数	Matlab	1stOpt
Exp1	a = 0.2318 b = 1.765 Goodness of fit: **SSE: 62.22** R-square: 0.9839 Adjusted R-square: 0.9828 RMSE: 2.037	**SSE: 62.2245194404254** RMSE: 1.91318212657258 R-Square: 0.985815218634677 Parameter Best Estimate --------- ------------- a 0.231771642275536 b 1.76457973001757
Exp2	a=1.824e+04 b = 0.9726 c=-1.824e+04 d = 0.9725	**SSE: 4.49583934179677** RMSE: 0.51425785142548 R-Square: 0.998899120898117 Parameter Best Estimate

	Goodness of fit: **SSE: 31.51** R-square: 0.9918 Adjusted R-square: 0.99 RMSE: 1.557	--------- a b c d	---------- 39.0863242050617 3.34855006651047 -39.0679731199091 3.34869099922272

对Exp1函数，二者结果一致，但对Exp2函数，Matlab的结果明显是错误的，其目标函数值SSE是1stOpt的7倍之多。

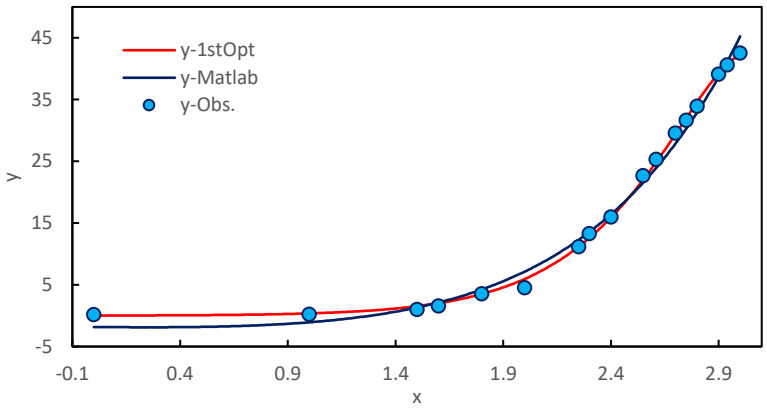

图 36.4 ”Exp2“指数函数计算结果对比图

36.3.2 Fourier：傅里叶函数

数据

x	2,5,10,20,30,50,100,200,400
y	5418.583,9479.431,14828.01,21052.6,23872.96,25784.02,26420.23,26445.85,26433.05

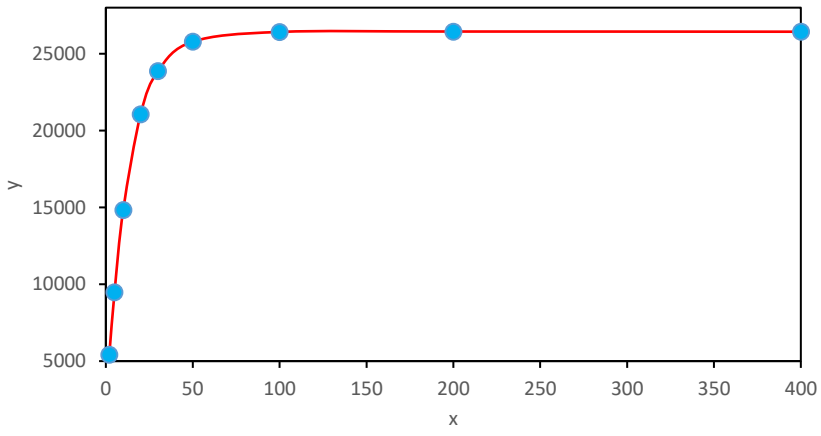

图 36.5 傅里叶函数（Fourier）验证数据图

Matlab内置Fourier函数项数最高为8，从1至8，每增加一项，参数数相应增加2个，本次验证数据数仅为9组，因此仅选取从1至3的傅里叶函数，保证参数数不大于数据数。

函数形式

Fourier1	$y = a_0 + a_1 \cdot \cos(x \cdot w) + b_1 \cdot \sin(x \cdot w)$
Fourier2	$y = a_0 + a_1 \cdot \cos(x \cdot w) + b_1 \cdot \sin(x \cdot w) + a_2 \cdot \cos(2 \cdot x \cdot w) + b_2 \cdot \sin(2 \cdot x \cdot w)$

Fourier3	$y = a_0 + a_1 \cdot \cos(x \cdot w) + b_1 \cdot \sin(x \cdot w) + a_2 \cdot \cos(2 \cdot x \cdot w) + b_2 \cdot \sin(2 \cdot x \cdot w) + a_3 \cdot \cos(3 \cdot x \cdot w) + b_3 \cdot \sin(3 \cdot x \cdot w)$

1stOpt代码：

```
Constant n=1;
Variable x, y;
Function y=a0+sum(i=1:n,a,b)(a*cos(i*x*w) + b*sin(i*x*w));
Data;
x=[2,5,10,20,30,50,100,200,400];
y=[5418.583,9479.431,14828.01,21052.6,23872.96,25784.02,26420.23,26445.85,26433.05];
```

Matlab与1stOpt计算结果

函数	Matlab	1stOpt
Fourier1	a0 = 2.114e+04 a1 = -8764 b1 = 9231 w = 0.01844 Goodness of fit: **SSE: 1.697e+08** R-square: 0.6761 Adjusted R-square: 0.4817	**SSE: 38625276.9394391** RMSE: 2071.64124252136 R-Square: 0.926253425659809 Parameter Best Estimate --------- ---------- a0 17537.8970408511 a1 9675.47591699509 w -17.2153756056715 b1 -3245.43987576037
Fourier2	a0= -1.352e+17 a1 = 1.802e+17 b1= -2.465e+14 a2 = -4.505e+16 b2 = 1.233e+14 w = -7.458e-06 Goodness of fit: **SSE: 1.378e+07** R-square: 0.9737 Adjusted R-square: 0.9298 RMSE: 2143	**SSE: 283281.135741205** RMSE: 177.413871354589 R-Square: 0.999459136218781 Parameter Best Estimate --------- ---------- a0 17236.509067496 a1 10764.832492643 w -1.63560631363068 b1 4077.55118922058 a2 -2810.24273368819 b2 -4517.42970148192
Fourier3	a0 = -4.228e+08 a1 = 5.789e+08 b1 = 2.888e+08 a2 = -1.671e+08 b2 = -2.171e+08 a3 = 1.1e+07 b3 = 4.865e+07 w = 0.003256 Goodness of fit: **SSE: 2530** R-square: 1 Adjusted R-square: 1 RMSE: 50.3	**SSE: 12.2178106713836** RMSE: 1.16513283321801 R-Square: 0.999999976672745 Parameter Best Estimate --------- ---------- a0 44608.8019944385 a1 44270.4508742796 w -0.984393862362967 b1 16026.6691906036 a2 64924.1941815572 b2 15556.6399208676 a3 35999.383590025 b3 -17469.4110693072

仅从计算结果看，Matlab对这三个内置拟合公式计算的结果都不正确，且与1stOpt计算结果SSE相比，误差均很大。

36.3.3 Gaussian：高斯函数

数据

x	0,20,30,50,60,70,80,100,120,130,140,150,160,180,200,240,320,360,400,440,480,520,560,600,640,680,72

	0,760,800,840,880,920,960,1000
y	4.4486,21.3705,39.9472,107.4536,155.4400,207.0610,254.0847,299.2430,254.0186,206.9618,155.3075,107.2876,68.1533,21.1025,4.1107,-0.9522,-1.0713,0.4890,7.9109,25.6786,44.6160,44.2607,24.5898,6.0160,-2.3448,-5.0594,-6.6304,-8.5346,-11.2359,-15.1607,-20.8160,-28.3768,-36.2845,-40.0000

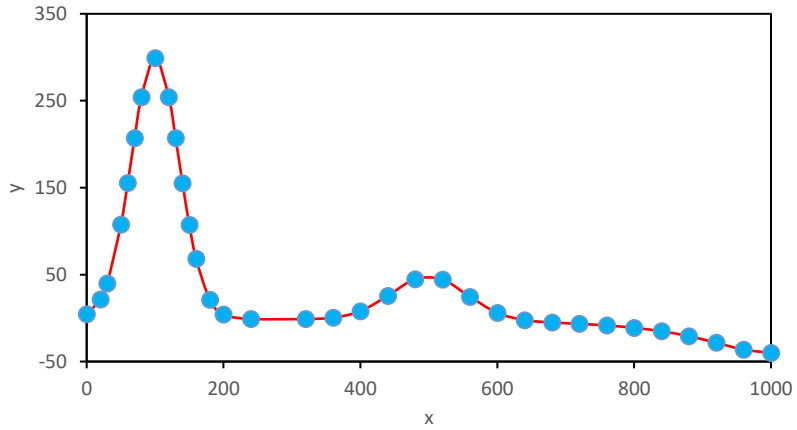

图 36.6 高斯函数（Gaussian）验证数据图

Matlab内置高斯函数项数最高为8，根据实际数据量及上图，仅选取从1至3项高斯函数。

Gaussian 函数

Gauss1	$y = a_1 \exp\left(-\left(\frac{x-b_1}{c_1}\right)^2\right)$
Gauss2	$y = a_1 \exp\left(-\left(\frac{x-b_1}{c_1}\right)^2\right) + a_2 \exp\left(-\left(\frac{x-b_2}{c_2}\right)^2\right)$
Gauss3	$y = a_1 \exp\left(-\left(\frac{x-b_1}{c_1}\right)^2\right) + a_2 \exp\left(-\left(\frac{x-b_2}{c_2}\right)^2\right) + a_3 \exp\left(-\left(\frac{x-b_3}{c_3}\right)^2\right)$

1stOpt代码：

```
Variable x, y;
Function y=a1*exp(-((x-b1)/c1)^2); //Gauss1
          //y=a1*exp(-((x-b1)/c1)^2)+a2*exp(-((x-b2)/c2)^2);    //Gauss2
          //y=a1*exp(-((x-b1)/c1)^2)+a2*exp(-((x-b2)/c2)^2)+a3*exp(-((x-b3)/c3)^2); //Gauss3
Data;
x=[0,20,30,50,60,70,80,100,120,130,140,150,160,180,200,240,320,360,400,440,480,520,560,600,640,680,720,
760,800,840,880,920,960,1000];
y=[4.4486,21.3705,39.9472,107.4536,155.4400,207.0610,254.0847,299.2430,254.0186,206.9618,155.3075,107
.2876,68.1533,21.1025,4.1107,-0.9522,-1.0713,0.4890,7.9109,25.6786,44.6160,44.2607,24.5898,6.0160,-
2.3448,-5.0594,-6.6304,-8.5346,-11.2359,-15.1607,-20.8160,-28.3768,-36.2845,-40.0000];
```

Matlab与1stOpt计算结果

函数	Matlab	1stOpt
Gauss1	a1 = 299.5 b1 = 99.99 c1 = 49.34 Goodness of fit: **SSE: 9975** R-square: 0.9654 Adjusted R-square: 0.9632 RMSE: 17.94	**SSE: 9975.36958039699** RMSE: 17.1287250825288 R-Square: 0.96551352609728 Parameter Best Estimate --------- ---------- a1 299.486768506001 b1 99.9855088360739 c1 49.3352095531223
Gauss2	a1 = 301.1 b1 = 100	**SSE: 4673.18298154511** RMSE: 11.7237604155296

	c1 = 49.6 a2 = -1.98 b2 = 116.8 c2 = 114.1 Goodness of fit: **SSE: 9973** R-square: 0.9654 Adjusted R-square: 0.9593 RMSE: 18.87	R-Square: 0.98865779978405 Parameter Best Estimate --------- ---------- a1 48.3777449619495 b1 498.571403273109 c1 72.3042633097725 a2 299.486768723683 b2 99.9855088238292 c2 49.3352094667924
Gauss3	a1 = 33.47 b1 = 100.7 c1 = 35.24 a2 = 185.4 b2 = 86.54 c2 = 44.89 a3 = 122.8 b3 = 121 c3 = 42.71 Goodness of fit: **SSE: 9974** R-square: 0.9654 Adjusted R-square: 0.9544 RMSE: 19.97	**SSE: 17.1184747071162** RMSE: 0.709566452285691 R-Square: 0.999944686437565 Parameter Best Estimate --------- ---------- a1 48.5618196506781 b1 499.777816877675 c1 -74.4440031336928 a2 299.486863499402 b2 99.9856655897989 c2 -49.3355190317666 a3 -71.3179308534723 b3 1249.86886164867 c3 338.184776997458

除了Gauss1外，Gauss2和Gauss3的Matlab结果均不正确，与1stOpt结果相比，相差很远。

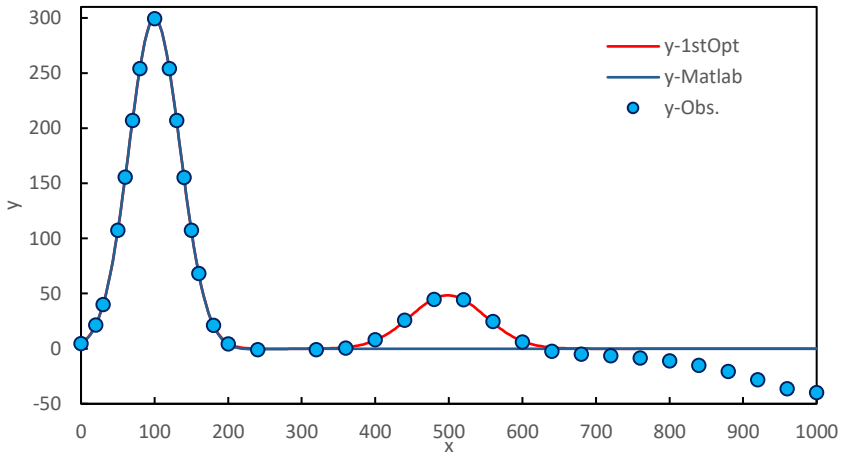

图 36.7 ” Gauss2“高斯函数计算结果对比图

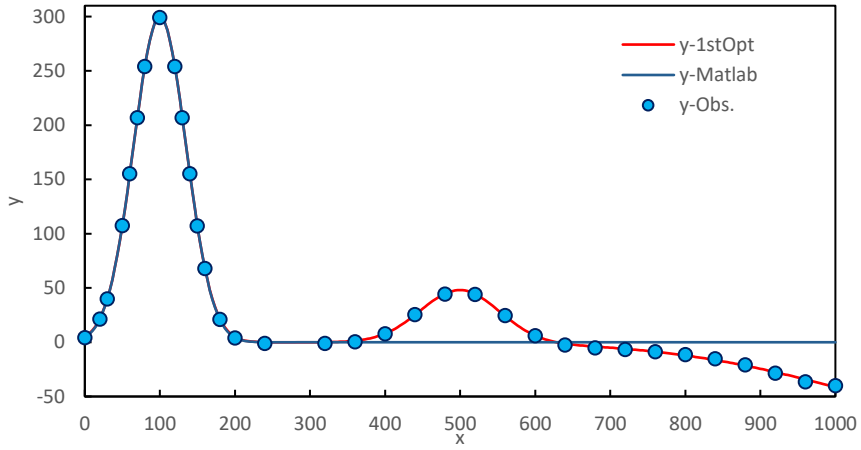

图 36.8 ” Gauss3“高斯函数计算结果对比图

36.3.4 Rational：有理函数

数据

x	0,33.333,66.667,100,133.333,166.667,200,300,400,500,600,700,800,900,1000
y	-57.51,121.746,167.57,180.572,170.651,131.428,89.633,32.923,9.106,2.714,0.79,0.188, 0.058,0.014,0.003

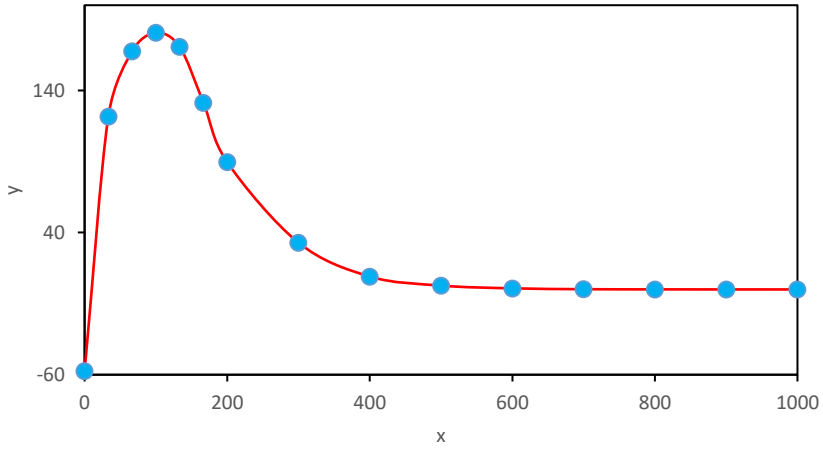

图 36.9 有理函数（Rational）验证数据图

Rational函数共有30种不同的拟合公式组合形式，选取以下6种。

Rational 函数

Rat01	$y = \frac{p_1}{x + q_1}$
Rat11	$y = \frac{p_1 x + p_2}{x + q_1}$
Rat22	$y = \frac{p_1 x^2 + p_2 x + p_3}{x^2 + q_1 x + q_2}$
Rat23	$y = \frac{p_1 x^2 + p_2 x + p_3}{x^3 + q_1 x^2 + q_2 x + q_3}$

Rat32	$y = \frac{p_1 x^3 + p_2 x^2 + p_3 x + p_4}{x^2 + q_1 x + q_2}$
Rat33	$y = \frac{p_1 x^3 + p_2 x^2 + p_3 x + p_4}{x^3 + q_1 x^2 + q_2 x + q_3}$

1stOpt代码：

```
Variable x,y;
Function y= (p1) / (x + q1); //Rat01
          //y=(p1*x + p2) / (x + q1); //Rat11
          //y= (p1*x^2 + p2*x + p3) / (x^2 + q1*x + q2);//Rat22
          //y=(p1*x^2 + p2*x + p3)/(x^3 + q1*x^2 + q2*x + q3);//Rat23
          //y=(p1*x^3 + p2*x^2 + p3*x + p4)/(x^2 + q1*x + q2);//Rat32
          //y=(p1*x^3 + p2*x^2 + p3*x + p4)/(x^3 + q1*x^2 + q2*x + q3);//Rat33
Data;
x=[0,33.333,66.667,100,133.333,166.667,200,300,400,500,600,700,800,900,1000];
y=[-57.51,121.746,167.57,180.572,170.651,131.428,89.633,32.923,9.106,2.714,0.79,0.188,0.058,0.014,0.003];
```

Matlab与1stOpt计算结果

函数	Matlab	1stOpt
Rat01	p1 = 2380 q1 = -19.51 Goodness of fit: **SSE: 8.542e+04** R-square: 0.009807 Adjusted R-square: -0.06636 RMSE: 81.06	**SSE: 65172.1078402097** RMSE: 65.9151514399179 R-Square: 0.286866571324611 Parameter Best Estimate --------- ---------- p1 38402.2295560357 q1 347.661227191482
Rat11	p1 = 48.4 p2 = 1246 q1 = -15.91 Goodness of fit: **SSE: 5.496e+04** R-square: 0.3628 Adjusted R-square: 0.2566 RMSE: 67.68	**SSE: 54518.0999431176** RMSE: 60.2871461939263 R-Square: 0.368021293230501 Parameter Best Estimate --------- ---------- p1 -509.951820225121 p2 404417.273100407 q1 3467.29400075593
Rat22	p1 = -41.47 p2 = 2.78e+04 p3 = -2.96e+05 q1 = -40.35 q2 = 5570 Goodness of fit: **SSE: 2470** R-square: 0.9714 Adjusted R-square: 0.9599 RMSE: 15.72	**SSE: 2470.07621283324** RMSE: 12.8324490071933 R-Square: 0.971366654886411 Parameter Best Estimate --------- ---------- p1 -41.4596543583183 p2 27796.3102692117 p3 -295977.362709882 q1 -40.3735552491385 q2 5568.98271999651
Rat23	p1 = 78.09 p2 = -103.9 p3 = -2.953 q1 = 24.68 q2 = -32.35 q3 = 0.05131 Goodness of fit: **SSE: 1.299e+05** R-square: -0.5062 Adjusted R-square: -1.343 RMSE: 120.2	**SSE: 50.4868900994394** RMSE: 1.83460967873168 R-Square: 0.999421606918053 Parameter Best Estimate --------- ---------- p1 -4099.33217955352 p2 2756823.21489415 p3 -6390666.04573042 q1 -211.740997416574 q2 22462.3211904212 q3 111115.959382505

Rat32	p1 = 0.128 p2 = -178.2 p3 = 5.818e+04 p4 = -5.246e+05 q1 = 6.561 q2 = 9445 Goodness of fit: **SSE: 898.7** R-square: 0.9896 Adjusted R-square: 0.9838 RMSE: 9.993	**SSE: 898.694103287714** RMSE: 7.74034496771177 R-Square: 0.989582257309595 Parameter Best Estimate --------- ---------- p1 0.128007447439966 p2 -178.154778485899 p3 58171.7603556022 p4 -524596.653325421 q1 6.54536629716628 q2 9444.80600308068
Rat33	p1 = 3.941 p2 = -7509 p3 = 3.41e+06 p4 = -1.076e+07 q1 = -214.7 q2 = 2.372e+04 q3 = 1.871e+05 Goodness of fit: **SSE: 41.37** R-square: 0.9995 Adjusted R-square: 0.9992 RMSE: 2.274	**SSE: 41.3669963700854** RMSE: 1.66066244553362 R-Square: 0.999520470065974 Parameter Best Estimate --------- ---------- p1 3.94079727739296 p2 -7509.26508469821 p3 3410306.5314588 p4 -10758807.5353447 q1 -214.653592277367 q2 23715.5772009434 q3 187068.210670941

上面6种函数中，Matlab结果只有一半的成功率，此外还需注意Matlab对后三种函数的计算结果不是稳定唯一的。

36.3.5 Sum of Sin：正弦和函数

数据

x	6800,6500,6350,5800,4900,4500,4000,3500,3000
y	-32.5,0.01,25,50,75,100,150,120,31

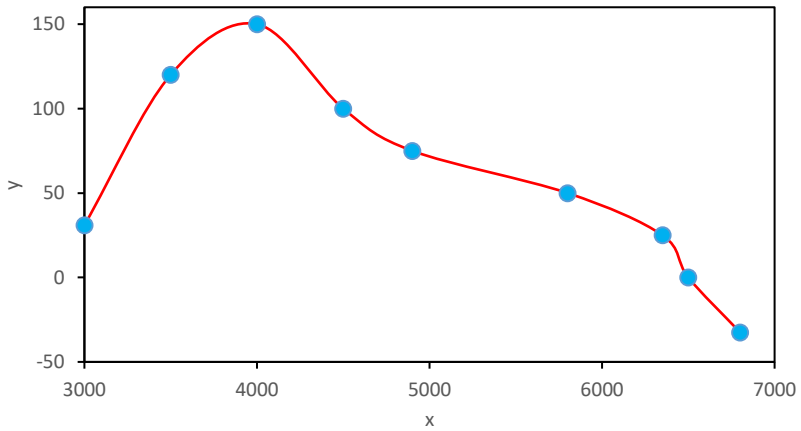

图 36.10 正弦和函数（Sum of Sin）验证数据图

Matlab内置Sum of Sin函数项数最高为8，根据实际数据量及上图，仅选取从1至3项正弦和函数。

Sum of Sin 函数

Sin1	$y = a_1 \cdot \sin(b_1 \cdot x + c_1)$
Sin2	$y = a_1 \cdot \sin(b_1 \cdot x + c_1) + a_2 \cdot \sin(b_2 \cdot x + c_2)$
Sin3	$y = a_1 \cdot \sin(b_1 \cdot x + c_1) + a_2 \cdot \sin(b_2 \cdot x + c_2) + a_3 \cdot \sin(b_3 \cdot x + c_3)$

1stOpt代码：

```
Constant n=1;    //n=1,2,3
Function y=Sum(n,a,b,c)(a*sin(b*x+c));
Data;
x=[6800,6500,6350,5800,4900,4500,4000,3500,3000];
y=[-32.5,0.01,25,50,75,100,150,120,31];
```

Matlab拟合工具箱与1stOpt计算结果

函数	Matlab	1stOpt
Sin1	a1 = 115.1 b1 = 0.0007424 c1 = -1.652 Goodness of fit: **SSE: 4613** R-square: 0.834 Adjusted R-square: 0.7787 RMSE: 27.73	**SSE: 4613.47367233302** RMSE: 22.6408521441835 R-Square: 0.834026943766448 Parameter Best Estimate --------- ---------- a1 115.088799943611 b1 218.02578779141 c1 -81148.8279322517
Sin2	a1 = 131.3 b1 = 0.0009764 c1 = -2.855 a2 = 52.43 b2 = 0.002129 c2 = 6.597 Goodness of fit: **SSE: 196.4** R-square: 0.9929 Adjusted R-square: 0.9812 RMSE: 8.091	**SSE: 15.0463415303111** RMSE: 1.29298713280145 R-Square: 0.999458663497979 Parameter Best Estimate --------- ---------- a1 146.398522272567 b1 0.000597128418062707 c1 -7.5484064889269 a2 -83.2946413909831 b2 -2.97270286953154 c2 -9.75816905890932
Sin3	a1 = 114.8 b1 = 0.00069 c1 = -1.246 a2 = 136.5 b2 = 0.003461 c2 = 2.55 a3 = 152.8 b3 = 0.00333 c3 = -6.121 Goodness of fit: **SSE: 65.02** R-square: 0.9977 Adjusted R-square: NaN RMSE: NaN	**SSE: 1.52638006364082E-19** RMSE: 1.30229714293237E-10 R-Square: 1 Parameter Best Estimate --------- ---------- a1 -17.8934285745309 b1 0.121901030220331 c1 26.305448755723 a2 130.843026332863 b2 -0.000652947630242742 c2 -8.13706797488529 a3 45.7927900482939 b3 0.0199949846230911 c3 -17.0205782929297

“Sum of Sin”函数因为含有周期性三角函数，即使目标函数值（SSE）一样，对应的参数组值也有可能是不同的，也即多解。从目标函数值看，Matlab给出的结果仅“Sin1”正确，其余两个均不对。

36.3.6 Weibull：温布尔函数

数据

x	1,1.1,1.2,1.3,1.4,1.5,1.6,1.7,1.8,1.9,2

y	-2.885,-4.463,-7.113,-11.799,-20.577,-38.093,-75.578,-162.239,-380.432,-983.866,-2833.698

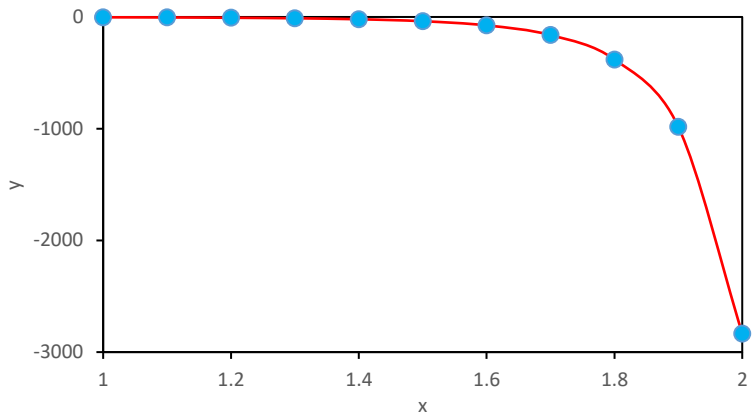

图 36.11 温布尔函数（Weibull）验证数据图

Weibull 函数：

$$y = a \cdot b \cdot x^{b-1} \cdot \exp(-a \cdot x^{b})$$

1stOpt代码：

```
Function y=a*b*x^(b-1)*exp(-a*x^b);
Data;
x=[1,1.1,1.2,1.3,1.4,1.5,1.6,1.7,1.8,1.9,2];
y=[-2.885,-4.463,-7.113,-11.799,-20.577,-38.093,-75.578,-162.239,-380.432,-983.866,-2833.698];
```

Matlab与1stOpt计算结果

Matlab	1stOpt
a = 1.912e-08	**SSE: 3.76470758625793E-7**
b = 1.132e-06	RMSE: 0.00018499895789038
Goodness of fit:	R-Square: 0.999999999999949
SSE: 9.177e+06	Parameter Best Estimate
R-square: -0.2539	--------- ----------
Adjusted R-square: -0.3932	a -0.499999987397168
RMSE: 1010	b 3.50000000402891

Matlab计算结果明显不对，其本身甚至也给出提示信息“Warning: A negative R-square is possible if the model does not contain a constant term and the fit is poor (worse than just fitting the mean). Try changing the model or using a different StartPoint.”，要求用户自己改变初值再试。

36.4 小结

测试了6大类18种内置函数，其中12种函数Matlab给出的结果都不正确，正确率仅为33.33%，作为科学数值计算的霸主，此结果应该是有些出乎意料的。说明了两点，一是尺有所短，寸有所长，相比1stOpt，Matlab在函数拟合方面还有很大的改进空间；二是在使用Matlab拟合工具箱的内置函数时，其自动赋初值的功能有很大缺陷，不可盲目相信并采纳其给出的结果，最好再使用类似1stOpt具有全局优化求解能力的软件再验证核实下，确保使用正确的结果。