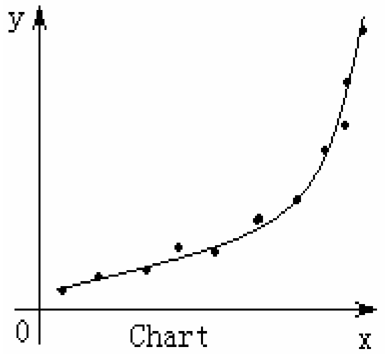
Extend Least Square Method

Yanping wang
Beijing yanshan hongta 17 - 2 – 602
Ww_yypp@163.com

Abstract. The Extend Least Square Method is a new method for data regression, which made the computation and results for data regression of one variable linear, multivariate linear and one variable nonlinear and multivariate nonlinear data calculating more easy and correct.

1. Principium

When studying the relating relations between the two variable numbers (x, y), we can get a series of binate data(x_1, y_1, x_2, y_2.....x_m, y_m), describing the data into the x - y orthogonal coordinate system(Chart), the points are found near a curve. suppose the one-variant non-linear variant of the curve such as (Formula 1)



$$y = a_0 + a_1 x_i^{\ k} \tag{1}$$

There into, a_0, a_1 and k are arbitrary real numbers

To set the curve variant, the numbers of a_0, a_1 and k, make the minimal value of the least square sum ($\sum(y_i - y)^2$) of the difference of the real observation value of y_i and the computing value of y using (Formula 1) as the optimization superior criterion.

Order:

$$\Phi = \sum (y_i - y)^2 = 0 \tag{2}$$

Take (Formula 1) to (Formula 2) to get :

$$\Phi = \sum \left(y_i - a_0 - a_1 x_i^{\ k} \right)^2 = 0 \tag{3}$$

When the square of $\sum(y_i - a_0 - a_1 x_i^k)$ is the smallest, we can use functionφ to get the partial differential coefficients of a_0, a_1 and k, make the three partial differential coefficients zero.

$$\frac{\partial \Phi}{\partial a_0} = -2\sum \left(y_i - a_0 - a_1 x_i^{\ k} \right) = 0 \tag{4}$$

$$\frac{\partial \Phi}{\partial a_1} = -2\sum x_i^{\ k} \left(y_i - a_0 - a_1 x_i^{\ k} \right) = 0 \tag{5}$$

$$\frac{\partial \Phi}{\partial k} = -2\sum x_i^{\ k} Ln(x_i) \left(y_i - a_0 - a_1 x_i^{\ k} \right) = 0 \tag{6}$$

Get three variant groups about a_0, a_1 and k which are the unknown numbers, edit the procedure for computer solve.

2. Example

Thirty suit data of three dimensions

Function(y)	Variable	Value of variable					
200	x_1	50	90	130	150	176	183
	x_2	2	17	93	180	400	500
300	x_1	62.1	100	140	180	230	284
	x_2	0.01	0.32	4.1	25	125	500
400	x_1	142	200	240	280	320	360
	x_2	0.01	1.12	8	38	128	340
500	x_1	0.01	0.1	0.45	2.7	11	32
	x_2	500	500	500	500	500	500
600	x_1	303	310	320	330	350	360
	x_2	0.01	0.011	0.04	0.08	0.35	0.65

Model of Extend Least Square Method regression

$$y = a_0 + a_1x_1^{k1} + a_2x_1^{k2} + a_3x_2^{k1} + a_4x_2^{k2} + a_5x_1^{k3}x_2^{k4} + a_6x_1^{k5}x_2^{k6} \tag{7}$$

Function in Formula 7:

y – Objective function
x_1 – Data of first dimensions
x_2 - Data of secondly dimensions

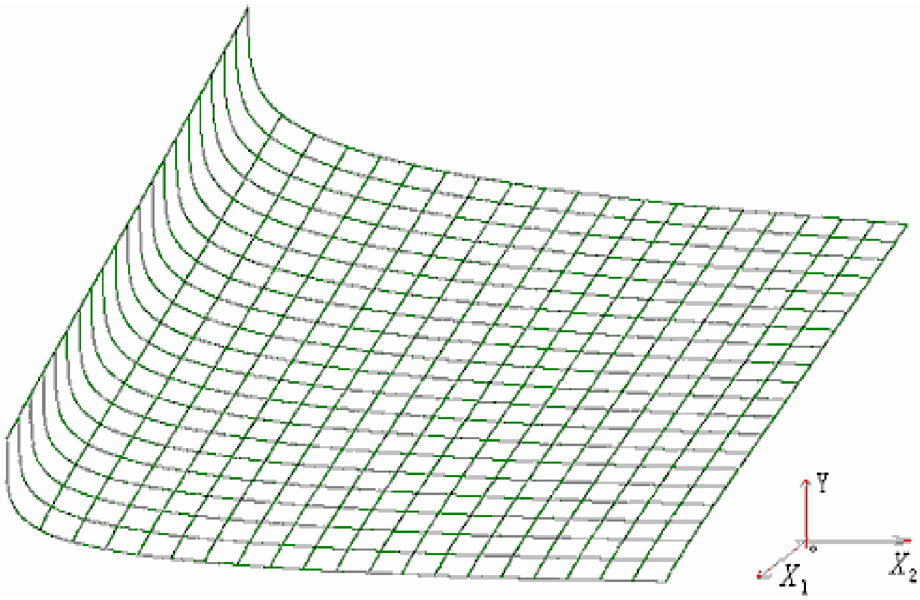
Coefficient and power of model

Coefficient of Model		Power of Model	
a_0	11.8311216825845	k_1	0.109999999998985
a_1	220.07825368487	k_2	1.17999999999898
a_2	0.322564255110857	k_3	0.649999999998985
a_3	-167.828014176716	k_4	0.319999999998985
a_4	0.0272258975662725	k_5	0.699999999998985
a_5	-0.14352770677907	k_6	1.05
a_6	-0.00118948794031524		

Test parameter of model

Parameter of test	Value
Correlation coefficient(R)	0.999949241084384
Biggest error	4.68829455350703
Equal error	1.104941
Equal relating error	0.0028556354582

First dimensions data is x_1 ,secondly dimensions is x_2 and objective function is y, curved face diagram of model



3. Conclusion

(1). Extend Least Square Method use the value of computing power to make the model function curves in different rates, to draw up the curves with different curve rates. It saves the complex ways of setting mechanism model and disposing linearization and makes the regression model and data drawing up better.

(2). Multiple nonlinear data regression in Extend Least Square Method , all variable and objective function, is at the same time the once regression mathematical model, since the regression , not only consider the contribution of objective function the variants take, and also consider the effects among the variants, so as to make the model correct.

(3). In the Extend Least Square Method Theory, the variant data can have many variants x^{k1}、 x^{k2} 、 … x^{kn}, i.e.(Formula 8). with the utility of the character, it can make the data more correct.

$$y = a_0 + a_1x^{k1} + a_2 x^{k2} + \ldots\ldots + a_nx^{kn} \quad \textbf{(8)}$$