第八章 非线性回归

§8.1 可化为线性回归的曲线回归

可线性化的曲线回归模型，也称为本质线性回归模型

$$y=\beta_0+\beta_1 e^x+\varepsilon \qquad (8.1)$$

只须令$x'=e^x$即可化为y对x'是线性的形式

$$y=\beta_0+\beta_1 x'+\varepsilon$$

需要指出的是，新引进的自变量只能依赖于原始变量，而不能与未知参数有关。

§8.1 可化为线性回归的曲线回归

$$y=\beta_0+\beta_1 x+\beta_2 x^2+\cdots+\beta_p x^p+\varepsilon \qquad (8.2)$$

令$x_1=x, x_2=x^2, \cdots, x_p=x^p$,

于是得到y关于$x_1, x_2, \cdots, x_p$的线性表达式

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p+\varepsilon$$

(8.2)式本来只有一个自变量x，是一元p次多项式回归，在线性化后，变为p元线性回归。

§8.1 可化为线性回归的曲线回归

可线性化的曲线回归模型，也称为本质线性回归模型

$$y=ae^{bx}e^{\varepsilon} \qquad (8.3)$$

对等式两边同时取自然对数，得：

$$\ln y=\ln a+bx+\varepsilon$$

令$y'=\ln y$，$\beta_0=\ln a$，$\beta_1=b$，

于是得到y′关于x的一元线性回归模型

$$y'=\beta_0+\beta_1 x+\varepsilon$$

§8.1 可化为线性回归的曲线回归

不可以线性化的曲线回归模型，也称为本质非线性回归模型

$$y= a\mathrm{e}^{bx}+\varepsilon \qquad (8.4)$$

当b未知时，不能通过对等式两边同时取自然对数的方法将回归模型线性化，只能用非线性最小二乘方法求解。

(8.3)式的误差项称为乘性误差项
(8.4)式的误差项称为加性误差项。
一个非线性回归模型是否可以线性化，不仅与回归函数的形式有关，而且与误差项的形式有关。

§8.1 可化为线性回归的曲线回归

在对非线性回归模型线性化时，总是假定误差项的形式就是能够使回归模型线性化的形式，为了方便，常常省去误差项，仅写出回归函数的形式。

例如把回归模型（8.3）式

$$y=ae^{bx}e^{\varepsilon}$$

简写为

$$y=ae^{bx}$$

§8.1 可化为线性回归的曲线回归

SPSS软件给出的10种常见的可线性化的曲线回归方程

英文名称	中文名称	方程形式
Linear	线性函数	$y=b_0+b_1 t$
Logarithm	对数函数	$y=b_0+b_1 \ln t$
Inverse	逆函数	$y=b_0+b_1/t$
Quadratic	二次曲线	$y=b_0+b_1 t+b_2 t^2$
Cubic	三次曲线	$y=b_0+b_1 t+b_2 t^2+b_3 t^3$
Power	幂函数	$y=b_0 t^{b_1}$
Compound	复合函数	$y=b_0 b_1^t$
S	S 型函数	$y=\exp(b_0+b_1/t)$
Logistic	逻辑函数	$y=\dfrac{1}{\frac{1}{u}+b_0 b_1^t}$ u 是预先给定的常数
Growth	增长曲线	$y=\exp(b_0+b_1 t)$
Exponent	指数函数	$y=b_0\exp(b_1 t)$

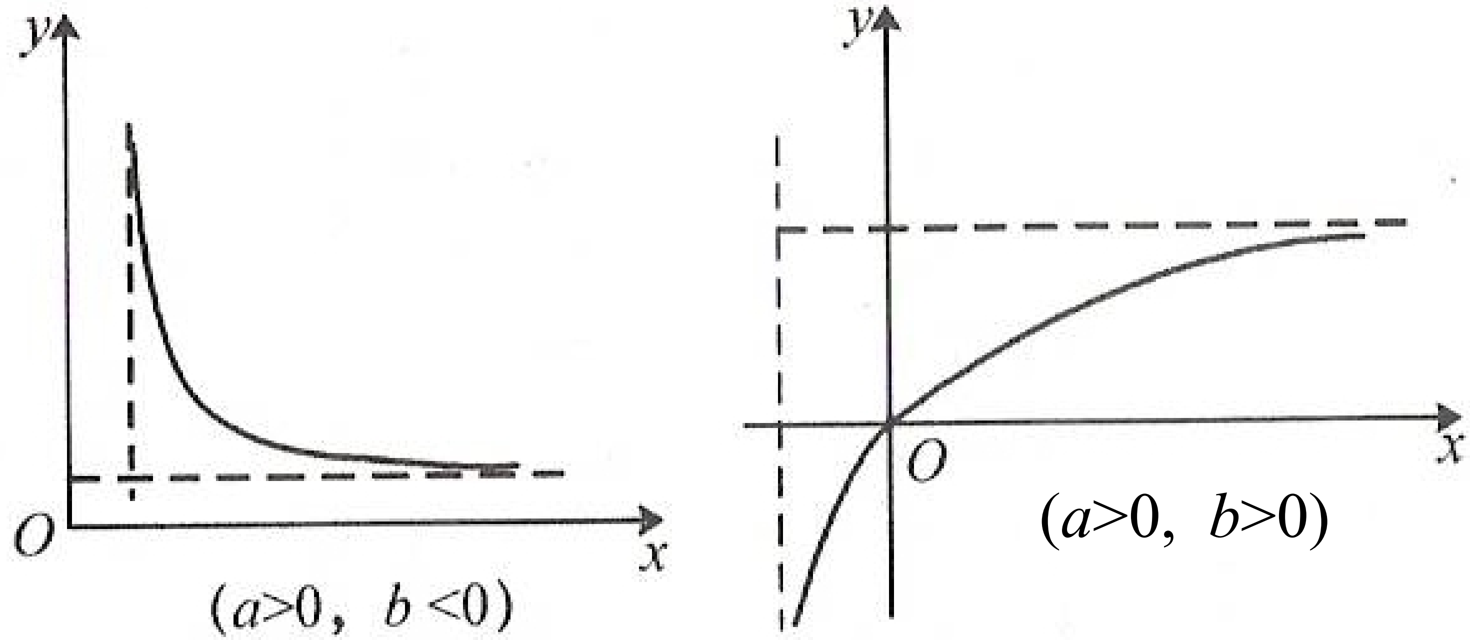
§8.1 可化为线性回归的曲线回归

除了以上SPSS软件中收入的几种曲线回归外，另外几种其他常用的曲线回归,例如

1. 双曲函数 $$y=\frac{x}{ax+b}$$

或等价地表示为 $$\frac{1}{y}=a+b\frac{1}{x}$$

§8.1 可化为线性回归的曲线回归



(a) 双曲函数

§8.1 可化为线性回归的曲线回归

2. S型曲线

$$y = \frac{1}{a + be^{-x}}$$

此S型曲线当a>0，b>0时，是x的增函数。

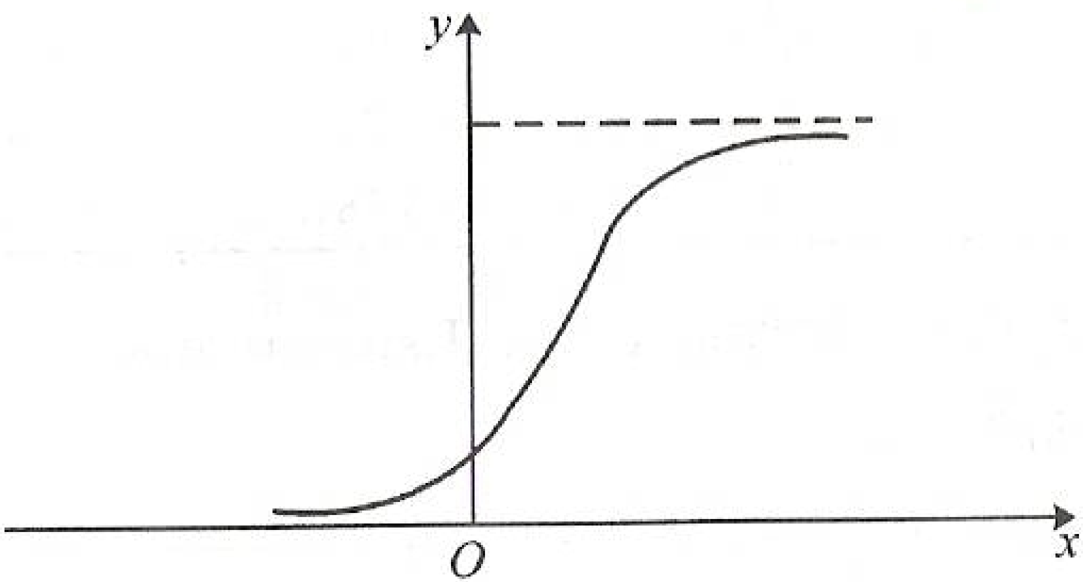
当x→+∞时，y→1/a； x→-∞时，y→0。

y=0与y=1/a是这条曲线的两条渐进线。

S型曲线有多种，其共同特点是曲线首先是缓慢增长，在达到某点后迅速增长，在超过某点后又变为缓慢增长，并且趋于一个稳定值。

S型曲线在社会经济等很多领域都有应用，例如某种产品的销售量与时间的关系，树木、农作物的生长与时间的关系等。

§8.1 可化为线性回归的曲线回归



(b) S形曲线

图 8.1

§8.1 可化为线性回归的曲线回归

SPSS软件中的S型曲线$y=\exp(b_0+b_1/t)$：

当$b_1<0$时是t的增函数，当t从右侧趋于0时，曲线趋于0；当$t\to+\infty$时，曲线以$y=\exp(b_0)$为渐进线，属于通常意义下的S型曲线。

当$b_1>0$时，曲线在t的正实轴上是t的减函数，不是通常意义下的S型曲线。

SPSS软件中的逻辑函数在$0<b_1<1$时也是S型曲线。

§8.1 可化为线性回归的曲线回归

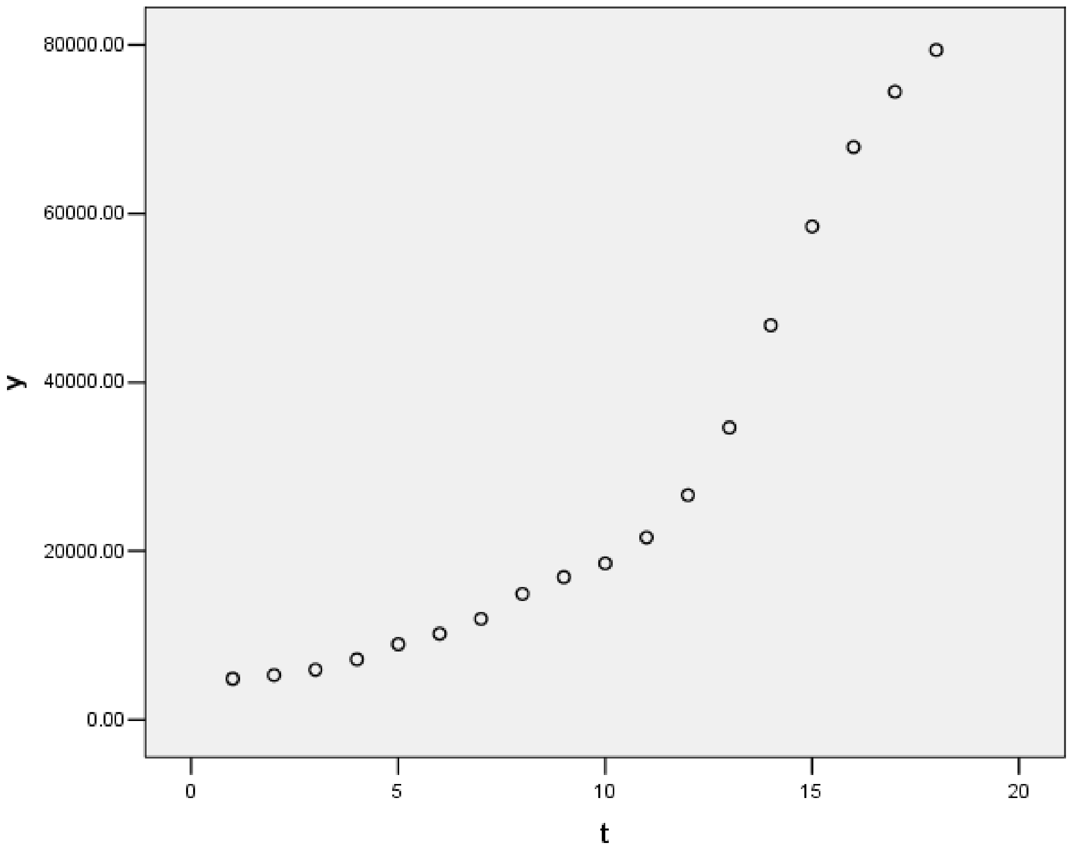
例8.1对GDP(国内生产总值)的拟合。我们选取GDP指标为因变量，单位为万亿元，拟合GDP关于时间t的趋势曲线。以1981年为基准年，取值为t=1，1998 年t=18，1981年至1998年的数据如表8.1。

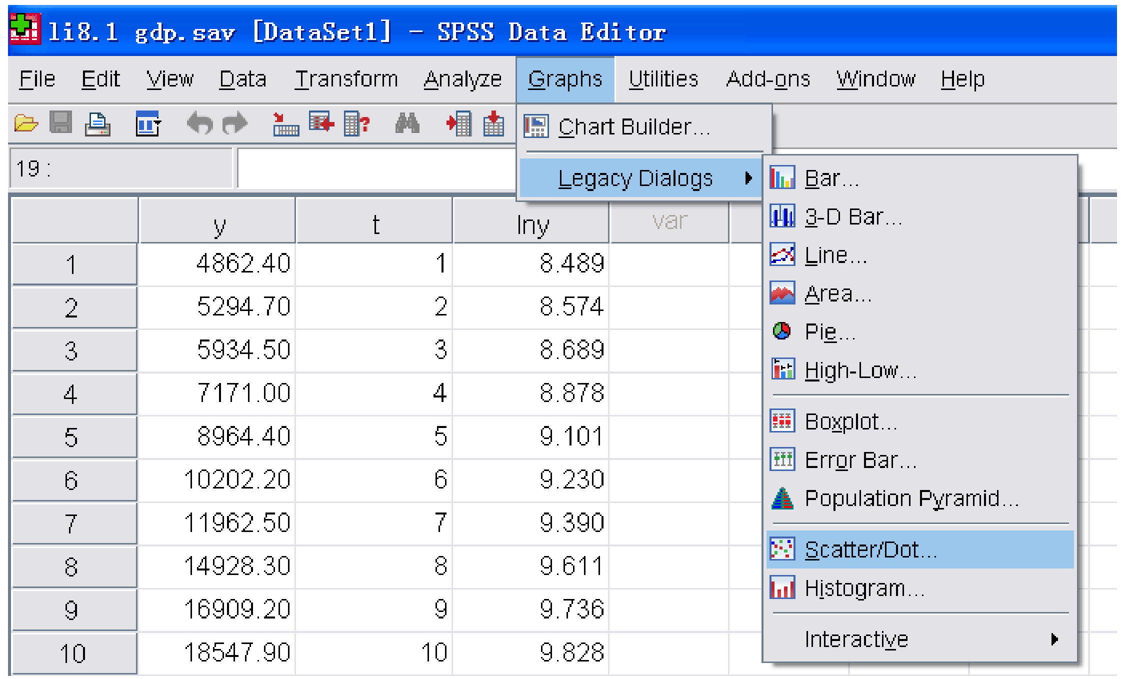
§8.1 可化为线性回归的曲线回归

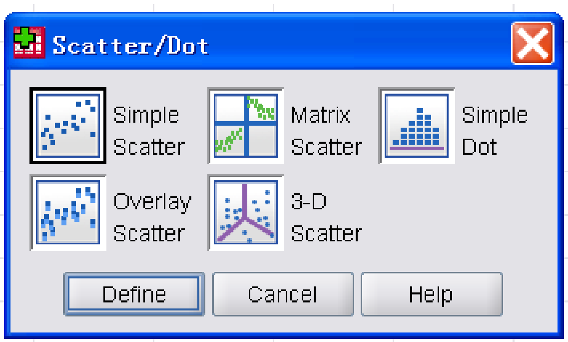
年份	t	y	$\hat{y}$	e_i	$y'=\ln y$
1981	1	4862.4	4296.35	566.05	8.489
1982	2	5294.7	5123.04	171.66	8.574
1983	3	5934.5	6108.80	-174.30	8.689
1984	4	7171.0	7284.24	-113.24	8.878
1985	5	8964.4	8685.86	278.54	9.101
1986	6	10202.2	10357.16	-154.96	9.230
1987	7	11962.5	12350.06	-387.56	9.390
1988	8	14928.3	14726.42	201.88	9.611
1989	9	16909.2	17560.04	-650.84	9.736
1990	10	18547.9	20938.89	-2390.99	9.828
1991	11	21617.8	24967.89	-3350.09	9.981
1992	12	26638.1	29772.14	-3134.04	10.190
1993	13	34634.4	35500.81	-866.41	10.453
1994	14	46759.4	42331.77	4427.63	10.753
1995	15	58478.1	50477.13	8000.97	10.976
1996	16	67884.6	60189.80	7694.80	11.126
1997	17	74462.6	71771.35	2691.25	11.218
1998	18	79395.7	85581.38	-6185.68	11.282

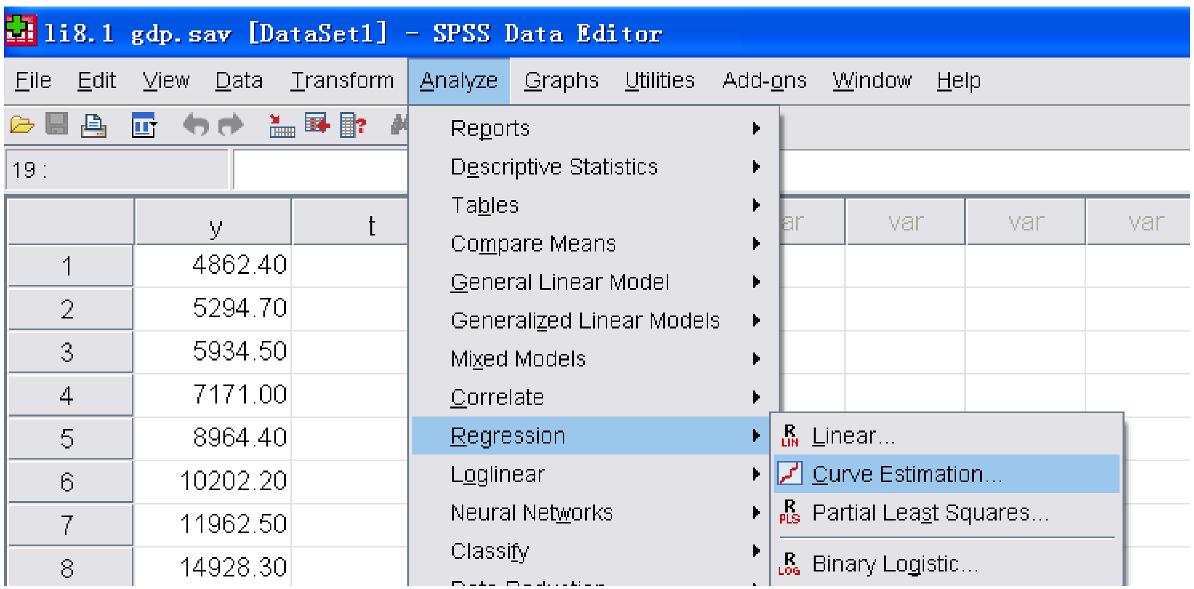
§8.1 可化为线性回归的曲线回归

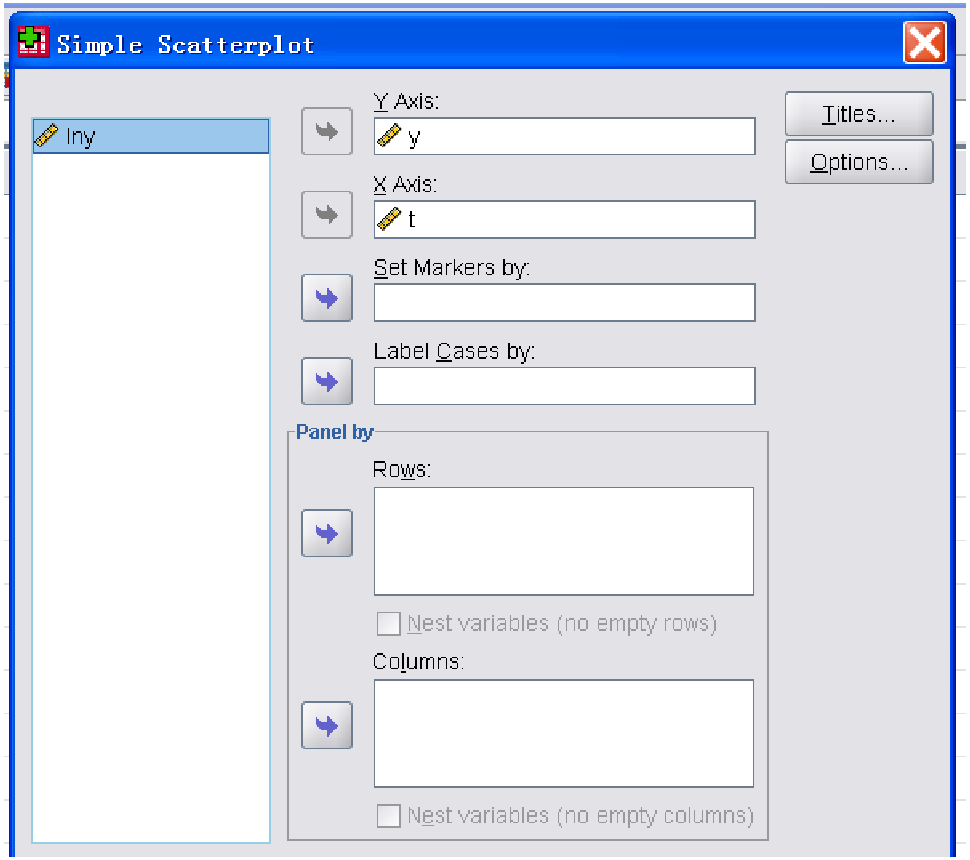
1. 直接用SPSS软件的Curve Estimation命令计算。首先画出GDP对时间的散点图，见图8.2。



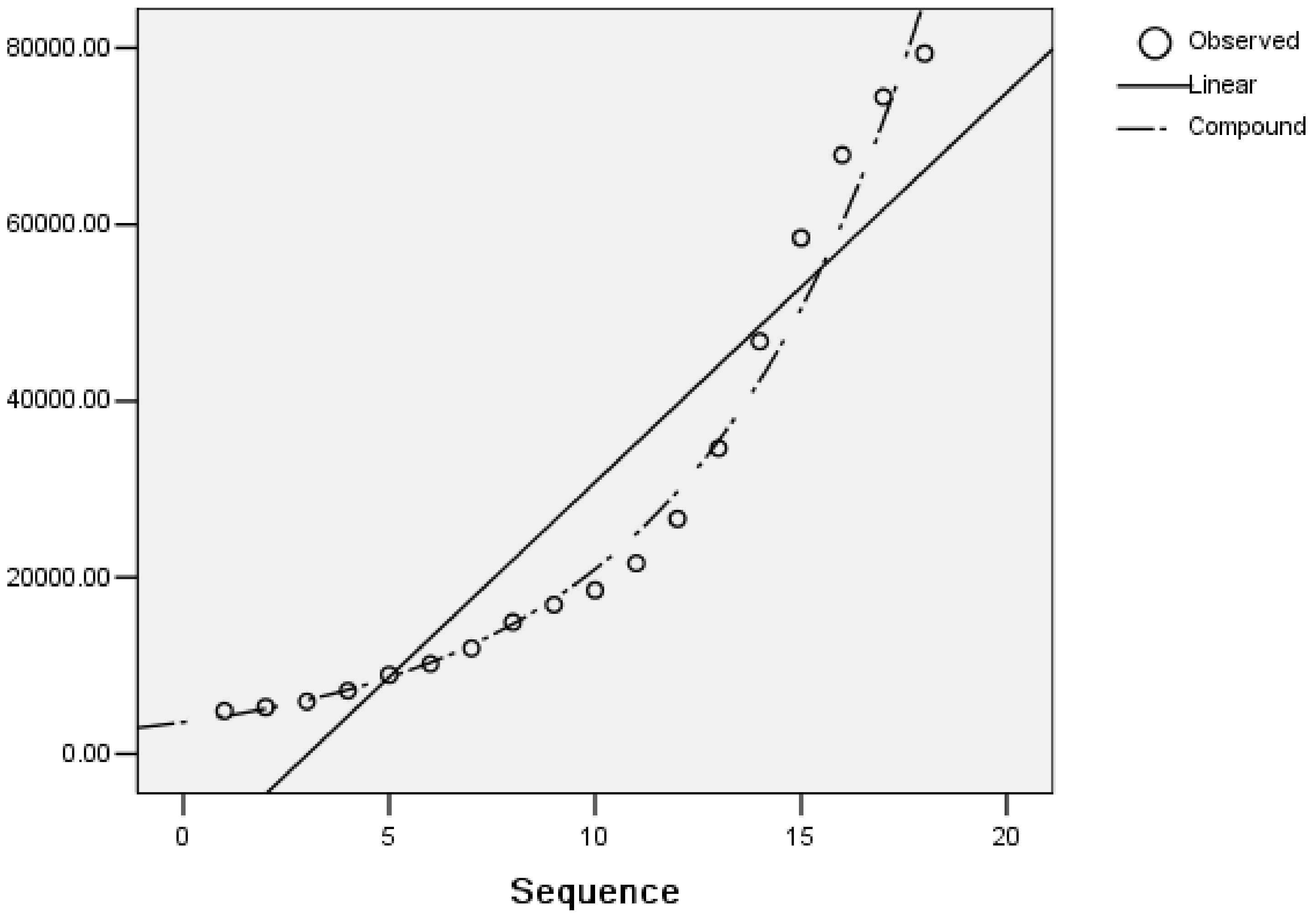

li8.1 gdp.sav [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
Chart Builder...
Legacy Dialogs
Bar...
3-D Bar...
Line...
Area...
Pie...
High-Low...
Boxplot...
Error Bar...
Population Pyramid...
Scatter/Dot...
Histogram...
Interactive
19 :
y t lny var
1 4862.40 1 8.489
2 5294.70 2 8.574
3 5934.50 3 8.689
4 7171.00 4 8.878
5 8964.40 5 9.101
6 10202.20 6 9.230
7 11962.50 7 9.390
8 14928.30 8 9.611
9 16909.20 9 9.736
10 18547.90 10 9.828


Scatter/Dot
Simple Scatter
Matrix Scatter
Simple Dot
Overlay Scatter
3-D Scatter
Define
Cancel
Help


li8.1 gdp.sav [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
19 :
y
t
var
var
var
1 4862.40
2 5294.70
3 5934.50
4 7171.00
5 8964.40
6 10202.20
7 11962.50
8 14928.30
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Linear...
Curve Estimation...
Partial Least Squares...
Binary Logistic...


Simple Scatterplot
lny
Y Axis:
y
X Axis:
t
Set Markers by:
Label Cases by:
Panel by
Rows:
Nest variables (no empty rows)
Columns:
Nest variables (no empty columns)
Titles...
Options...

§8.1 可化为线性回归的曲线回归



§8.1 可化为线性回归的曲线回归

表 8. 2　　　　　　线性回归 $y=b_0+b_1t$

Multiple R	.92528
R Square	.85615
Adjusted R Square	.84716
Standard Error	9964.23063

Analysis of Variance:

	DF	Sum of Squares	Mean Square	F	Signif F
Regression	1	9454779005.1	9454779005.1	95.22782	.0000
Residuals	16	1588574273.6	99285892.1		

Variable	B	SE B	Beta	T	Sig T
Time	4417.522807	452.685809	.925284	9.758	.0000
(Constant)	-13374.922222	4900.032018		-2.730	.0148

§8.1 可化为线性回归的曲线回归

表8.3

复合函数回归 $y=b_0\ b_1^t$

Multiple R	.99593
R Square	.99188
Adjusted R Square	.99138
Standard Error	.08760

Analysis of Variance:

	DF	Sum of Squares	Mean Square	F	Signif F
Regression	1	15.004878	15.004878	1955.31315	.0000
Residuals	16	.122782	.007674		

Variable	B	SE B	Beta	T	Sig T
Time	1.192417	.004746	2.707250	251.269	.0000
(Constant)	3603.061130	155.215413		23.213	.0000

§8.1 可化为线性回归的曲线回归

为了与线性回归的拟合效果直接相比，可以先储存复合函数回归的残差序列，然后计算出

复合函数回归的 SSE =262467769=2.625×10^8，

R^2=1-262467769/11043353279=0.97623，

拟合效果明显优于线性回归，当然应该采用复合函数回归。

§8.1 可化为线性回归的曲线回归

复合函数回归b_0=3603.06,等比系数b_1=1.192417，回归方程为

$$\hat{y} = 3603.06(1.192417)^t$$

其中b_1=1.192417=119.2417%表示GDP的平均发展速度，平均增长速度为19.2417% 。

这里GDP是用的当年现价，在实际工作中可以用不变价格代替现价；对误差项的自相关做相应的处理；考虑到GDP的年增长速度会有减缓趋势，可以对回归函数增加适当的阻尼因子等改进方法。

§8.1 可化为线性回归的曲线回归

2. 线性化求解法。

对复合函数$y=b_0$两端取自然对数，得

$$\ln y=\ln b_0+\ln(b_1)\,t$$

令$y'=\ln y$，$\beta_0=\ln b_0$，$\beta_1=\ln(b_1)$，

于是得到y'关于t的线性回归方程

$$y'=\beta_0+\beta_1 t$$

计算出$y'=\ln y$的值列在表8.4中，用y'对t做一元线性回归，输出结果为：

§8.1 可化为线性回归的曲线回归

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.996[a]	.992	.991	8.7601E-02	.616

a. Predictors: (Constant), T

b. Dependent Variable: LNY

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	15.005	1	15.005	1955.313	.000
	Residual	.123	16	7.674E-03		
	Total	15.128	17			

§8.1 可化为线性回归的曲线回归

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	8.190	.043		190.106	.000
	T	.176	.004	.996	44.219	.000

其中$\hat{\beta}_0$=8.190，$\hat{\beta}_1$=0.176，得

$$\hat{b}_0 = e^{8.190} = 3604.7,\ \hat{b}_1 = e^{0.176} = 1.1924$$

与直接用 SPSS 软件的 Curve Estimation 命令计算的结果相一致。

§8.2 多项式回归

一、几种常见的多项式回归模型

一元二次多项式模型$y_i=\beta_0+\beta_1x_i+\beta_{11}+\varepsilon_i$

的回归函数$y_i=\beta_0+\beta_1x_i+\beta_{11}$是一条抛物线方程，通常称为二项式回归函数。

回归系数β_1为线性效应系数，β_{11}为二次效应系数。

相应地，回归模型 $y_i=\beta_0+\beta_1x_i+\beta_{11}+\beta_{111}+\varepsilon_i$

称为一元三次多项式模型。

§8.2 多项式回归

称回归模型

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{11} x_{i1}^2 + \beta_{22} x_{i2}^2 + \beta_{12} x_{i1} x_{i2} + \varepsilon_i$$

为二元二阶多项式回归模型。

它的回归系数中分别含有两个自变量的线性项系数 β_1 和 β_2，二次项系数 β_{11} 和 β_{22}，并含有交叉乘积项系数 β_{12}。

交叉乘积项表示 x_1 与 x_2 的交互作用。

§8.2 多项式回归

二、一个应用例子

例8.2 表8.5列出的数据是关于18个35岁～44岁经理的:

前两年平均年收入 x_1（千美元）

风险反感度 x_2

人寿保险额 y（千美元）

风险反感度是根据发给每个经理的标准调查表估算得到的；它的数值越大，风险反感就越厉害。

§8.2 多项式回归

研究人员想研究给定年龄组内的经理年平均收入，风险反感度和人寿保险的关系。研究者预计，在经理的收入和人寿保险额之间成立着二次关系，并有把握认为风险反感度对人寿保险额只有线性效应，而没有二次效应。但是，研究者对两个自变量是否对人寿保险额有交互效应，心中没底。因此，研究者拟合了一个二阶多项式回归模型

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{11} x_{i1}^2 + \beta_{22} x_{i2}^2 + \beta_{12} x_{i1} x_{i2} + \varepsilon_i$$

并打算先检验是否有交互效应，然后检验风险反感的二次效应。

§8.2 多项式回归

序号	x_{i1}	x_{i2}	y_i
1	66.290	7	196
2	40.964	5	63
3	72.996	10	252
4	45.010	6	84
5	57.204	4	126
6	26.852	5	14
7	38.122	4	49
8	35.840	6	49
9	75.796	9	266
10	37.408	5	49
11	54.376	2	105
12	46.186	7	98
13	46.130	4	77
14	30.366	3	14
15	39.060	5	56
16	79.380	1	245
17	52.766	8	133
18	55.916	6	133

§8.2 多项式回归

回归采用逐个引入自变量的方式，

依次引入自变量 x_1、x_2、 x_1^2 x_2^2 、x_1x_2，方法如下：

在线性回归对话框中，点入 y 与 x1，然后点 Block 1 of Next,

这时自变量框变为空白，再把 x_1、x_2 同时点入自变量框中，

然后再点 Block 2 of Next, 自变量框又变为空白，

再把 x_1、x_2、

x_1^2

§8.2 多项式回归

ANOVA[f]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	104474.1	1	104474.107	468.471	.000[a]
	Residual	3568.170	16	223.011		
	Total	108042.3	17			
2	Regression	106758.4	2	53379.192	623.641	.000[b]
	Residual	1283.893	15	85.593		
	Total	108042.3	17			
3	Regression	107996.8	3	35998.917	11070.294	.000[c]
	Residual	45.526	14	3.252		
	Total	108042.3	17			
4	Regression	107999.9	4	26999.964	8274.003	.000[d]
	Residual	42.422	13	3.263		
	Total	108042.3	17			
5	Regression	108005.8	5	21601.164	7110.202	.000[e]
	Residual	36.457	12	3.038		
	Total	108042.3	17			

a. Predictors: (Constant), x1

b. Predictors: (Constant), x1, x2

c. Predictors: (Constant), x1, x2, x11

d. Predictors: (Constant), x1, x2, x11, x22

e. Predictors: (Constant), x1, x2, x11, x22, x12

f. Dependent Variable: y

§8.2 多项式回归

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part
1	(Constant)	-140.550	12.170		-11.548	.000			
	x1	5.040	.233	.983	21.644	.000	.983	.983	.983
2	(Constant)	-158.768	8.324		-19.074	.000			
	x1	4.843	.149	.945	32.472	.000	.983	.993	.914
	x2	5.201	1.007	.150	5.166	.000	.391	.800	.145
3	(Constant)	-62.349	5.200		-11.989	.000			
	x1	.840	.207	.164	4.052	.001	.983	.735	.022
	x2	5.685	.198	.164	28.738	.000	.391	.992	.158
	x11	.037	.002	.785	19.515	.000	.986	.982	.107
4	(Constant)	-60.910	5.414		-11.250	.000			
	x1	.930	.227	.182	4.090	.001	.983	.750	.022
	x2	4.453	1.278	.129	3.483	.004	.391	.695	.019
	x11	.036	.002	.760	15.815	.000	.986	.975	.087
	x22	.116	.119	.038	.975	.347	.565	.261	.005
5	(Constant)	-65.386	6.123		-10.679	.000			
	x1	1.017	.228	.198	4.460	.001	.983	.790	.024
	x2	5.217	1.349	.151	3.868	.002	.391	.745	.021
	x11	.036	.002	.758	16.342	.000	.986	.978	.087
	x22	.166	.120	.055	1.383	.192	.565	.371	.007
	x12	-.020	.014	-.046	-1.401	.186	.707	-.375	-.007

a. Dependent Variable: y

§8.2 多项式回归

表8.6

变量	偏平方和	残差	检验系数	偏 F 值
X_1	104474	3567	β_1	-
$X_2 \mid X_1$	2284	1283	β_2	-
$X_1^2 \mid X_1, X_2$	1238	45	β_{11}	1238/（45/14）=385
$X_2^2 \mid X_1, X_2, X_1^2$	3	42	β_{22}	3/（42/13）=0.93
$X_1X_2 \mid X_1, X_2, X_1^2, X_2^2$	6	36	β_{12}	6/（36/12）=2.00
合计	108005	5		

§8.2 多项式回归

得最终的回归方程为：

$$\hat{y} = -62.349 + 0.840x_1 + 5.685x_2 + 0.0371 x_1^2$$

括号中的数值是标准化回归系数。

这样，研究者就可用这个回归方程来进一步研究经理的年平均收入和风险反感对人寿保险额的效应。从标准化回归系数看到，年平均收入的二次效应对人寿保险额的影响程度最大。

§8.2 多项式回归

【例8.3】 维生素C注射液因长期放置会渐变成微黄色，中国药典规定可以用焦亚硫酸钠等作为抗氧剂。本实验考虑3个因素，分别是

EDTA（X_1）

无水碳酸钠（X_2）

焦亚硫酸钠（X_3）

每个因素各取7个水平，选用U_7（7^4）均匀设计表，取其中的第1、2、3列，实验安排与结果见表6.9。

§8.2 多项式回归

表6.9 实验设计与结果

实验号	EDTA X_1（g）	无水碳酸钠 X_2（g）	焦亚硫钠 X_3（g）	吸收度 y	$1/y$
1	0.00	30	0.6	1.160	0.862
2	0.02	38	1.2	0.312	3.205
3	0.04	46	0.4	0.306	3.263
4	0.06	26	1.0	1.318	0.759
5	0.08	34	0.2	0.877	1.140
6	0.10	42	0.8	0.147	6.803
7	0.12	50	1.4	0.204	4.902

§8.2 多项式回归

首先做线性回归，回归的计算程序参照例6.1，得回归方程

$$y = 2.63 + 0.77\ X_1 - 0.0524\ X_2 - 0.087\ X_3$$

回归模型的*P*值=0.1040；

决定系数（*R-square*）= 83.9% ；

调整的决定系数（*AdjR-sq*）= 67.8%。

可见线性回归的效果不够好，以下使用二次多项式回归。

§8.2 多项式回归

使用逐步回归，回归方程的具体形式是：

$$y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + B_{11}X_1^2 + B_{22}X_2^2 + B_{33}X_3^2$$
$$+ B_{12}X_1X_2 + B_{13}X_1X_3 + B_{23}X_2X_3$$

做变量替换转化为9个自变量的线性回归。

$$X_{11} = X_1^2,\ X_{22} = X_2^2,\ X_{33} = X_3^2$$
$$X_{12} = X_1X_2,\ X_{13} = X_1X_3,\ X_{23} = X_2X_3$$

§8.2 多项式回归

表6.10　回归变量表

X_1	X_2	X_3	X_{11}	X_{22}	X_{33}	X_{12}	X_{13}	X_{23}	y
0.00	30	0.6	0.0000	900	0.360	0.00	0.000	18.0	1.160
0.02	38	1.2	0.0004	1444	1.440	0.76	0.024	45.6	0.312
0.04	46	0.4	0.0016	2116	0.160	1.84	0.016	18.4	0.306
0.06	26	1.0	0.0036	676	1.000	1.56	0.060	26.0	1.318
0.08	34	0.2	0.0064	1156	0.040	2.72	0.016	6.8	0.877
0.10	42	0.8	0.0100	1764	0.640	4.20	0.080	33.6	0.147
0.12	50	1.4	0.0144	2500	1.960	6.00	0.168	70.0	0.204

§8.2 多项式回归

这个线性回归只有7组观测数据却有10个未知参数，需要使用逐步回归逐个引入变量。

在SPSS软件逐步回归模块默认的进入变量P值=0.05，剔除变量P值=0.10的条件下，逐步回归只进行了一步就结束了，只选入了自变量x_2。为了更全面地了解回归的效果，可以把进入变量的条件放宽一些。

用Option选项把进入变量P值改为0.30，剔除变量P值改为0.50，重新做逐步回归。

§8.2 多项式回归

表6.12　逐步回归的输出结果（2）

Step	1	2	3	4	5
Constant	2.579	5.957	7.311	7.873	9.165
X_2 *Prob>F*	-0.0516 0.004	-0.2376 0.053	-0.3034 0.021	-0.3126 0.030	-0.378 0.016
X_{22} *Prob>F*		0.00245 0.100	0.00336 0.033	0.00323 0.048	0.0046 0.019
X_3 *Prob>F*			-0.292 0.107	-1.115 0.168	-1.430 0.033
X_{23} *Prob>F*				0.0206 0.251	0.0317 0.039
X_{13} *Prob>F*					-2.33 0.058
R-square	83.14	92.12	97.11	98.73	99.99

§8.2 多项式回归

此时的逐步回归共进行了5步，依次选入了X_2， $X_{22}=X_2^2$，X_3，$X_{23}=X_2X_3$，$X_{13}=X_1X_3$共5个变量，共计算出5个回归模型：

第一个回归模型最先选入的是X_2，说明无水碳酸钠的含量是最重要的影响因素；

第二个回归模型再选入的是$X_{22}=X_2^2$，进一步说明无水碳酸钠的含量是最重要的影响因素，并且说明y与X_2的关系是非线性的

$$y = 5.975 - 0.2375X_2 + 0.00245X_2^2$$

容易求出此方程在X_2=48.5≈48时达极小值y=0.197，比第6号实验值y=0.147略高。

§8.2 多项式回归

再看第三个回归方程：

$$y = 7.311 - 0.303X_2 + 0.00336X_2^2 - 0.29X_3$$

为使y值最小，X_3应该最大，取X_3=1.4，X_2的取值与X_3无关，容易求出此方程在X_2=45.1≈45，X_3=1.4时达极小值y=0.074，低于第6号实验值y=0.147。

§8.2 多项式回归

第四个回归方程是:

$$y = 7.873 - 0.3126X_2 + 0.00323X_2^2 - 1.115X_3 + 0.0206X_2X_3$$

在回归方程含有X_3的两项－1.115 X_3+0.0206 X_2X_3中，当X_2≤54时是X_3的减函数，根据对第二和第三两个回归方程的分析，两个方程中X_2的最优解分别是48和45，所以有理由认为X_2≤54，y是X_3的减函数，X_3越大y越小，因此取X_3=1.4。

把X_3=1.4代入以上方程中，解得X_2的极小值是X_2=43.9≈44，所以第四个回归方程的最优组合是X_2=44，X_3=1.4，此时最优预测值y=0.080，与第三个回归方程的最优解基本相同。

§8.2 多项式回归

第五个方程是：

$$y = 9.16 - 0.379X_2 + 0.00406\,X_2^2 - 1.43\,X_3 + 0.0317\,X_2X_3 - 2.33\,X_1X_3$$

其中包含了变量X_1，并且是作为与X_3的交互作用形式出现，说明EDTA对实验指标本身没有影响，只是通过焦亚硫酸钠对实验产生弱的影响。仿照对第四个回归方程求最优解的方法，首先确定X_1和X_3是y的减函数，分别取最大值X_1=0.12和X_3=1.4，然后再解得X_2=41.2≈41。最优预测值

y= －0.128<0，可以视为接近0。