Southwest Petroleum University
西南石油大学

第五章 回归分析

§5.3 一元非线性回归

理学院

- 预习：
  - 变量之间的相关性，假若不是直线或者线性关系，该如何确定其回归方程？
  - 确定非直线回归方程该用什么方法呢？

§5.3 一元非线性回归

曲线回归分析：是通过两个相关变量x与y的实际观测数据建立曲线回归方程，以揭示x与y间的曲线联系的形式。

曲线回归分析最困难和首要的工作是确定变量Y与x间的曲线关系的类型。通常通过两个途径来确定：

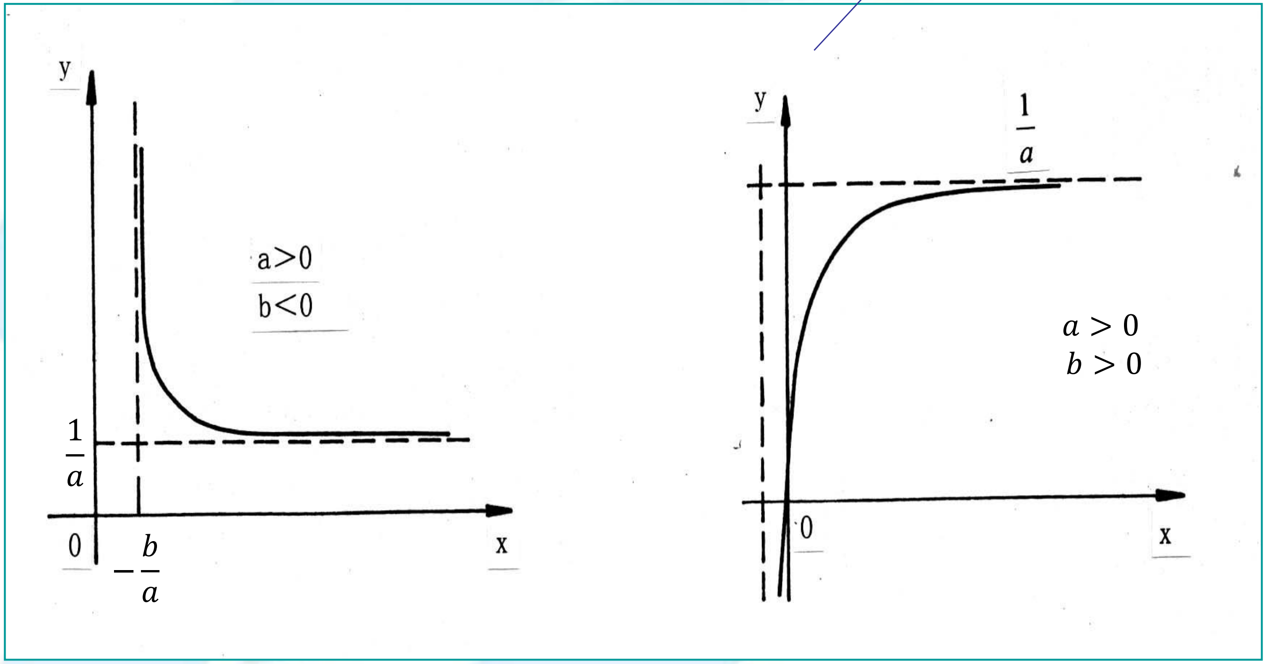
- 1. 利用有关的专业知识，根据已知的理论规律和实践经验。
- 2. 在没有已知的理论规律和经验可资利用时，则可用描点法将实测点在直角坐标纸上描出，观察实测点的分布趋势与哪一类已知的函数曲线最接近，然后再选用该函数关系式来拟合实测点。

可线性化的曲线函数类型

（1） 双曲线型

方法：变量替换

$$\frac{1}{y}=a+\frac{b}{x}$$

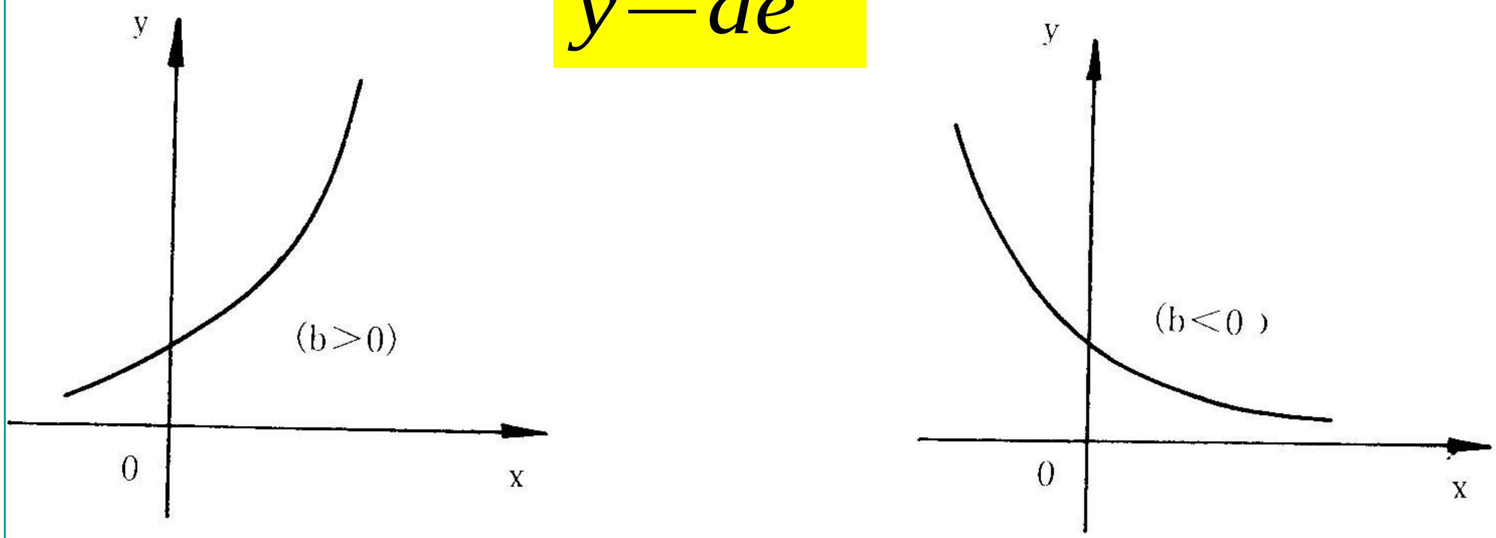


$$\frac{1}{y}=a+\frac{b}{x}$$

令 $u=\frac{1}{y}, v=\frac{1}{x}$ 得到 $u=a+bv$

（2） 指数曲线型

$$y = ae^{bx}$$



令$v = \ln y$，得到：

$$v = \ln a + bx$$

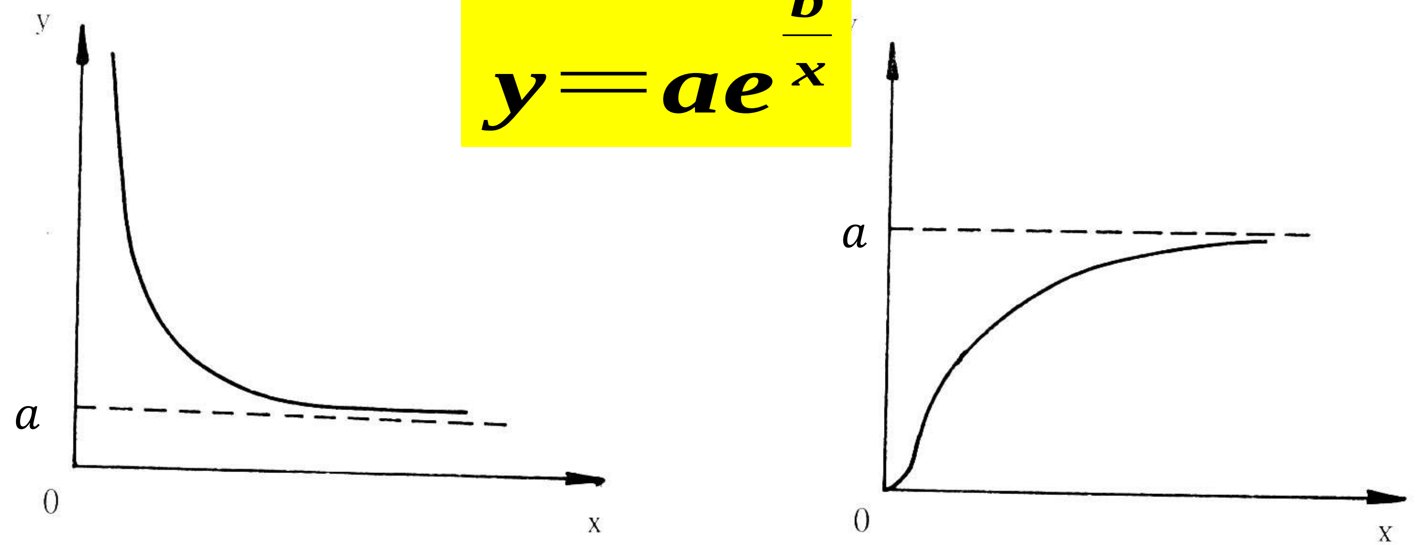
倒指数曲线型

$$y = ae^{\frac{b}{x}}$$

令 $y' = \ln y$，$x' = 1/x$，得到：

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

$$y = ae^{\frac{b}{x}}$$



$y = ae^{bx}$ （b>0）　　$y = ae^{b/x}$ （b<0）

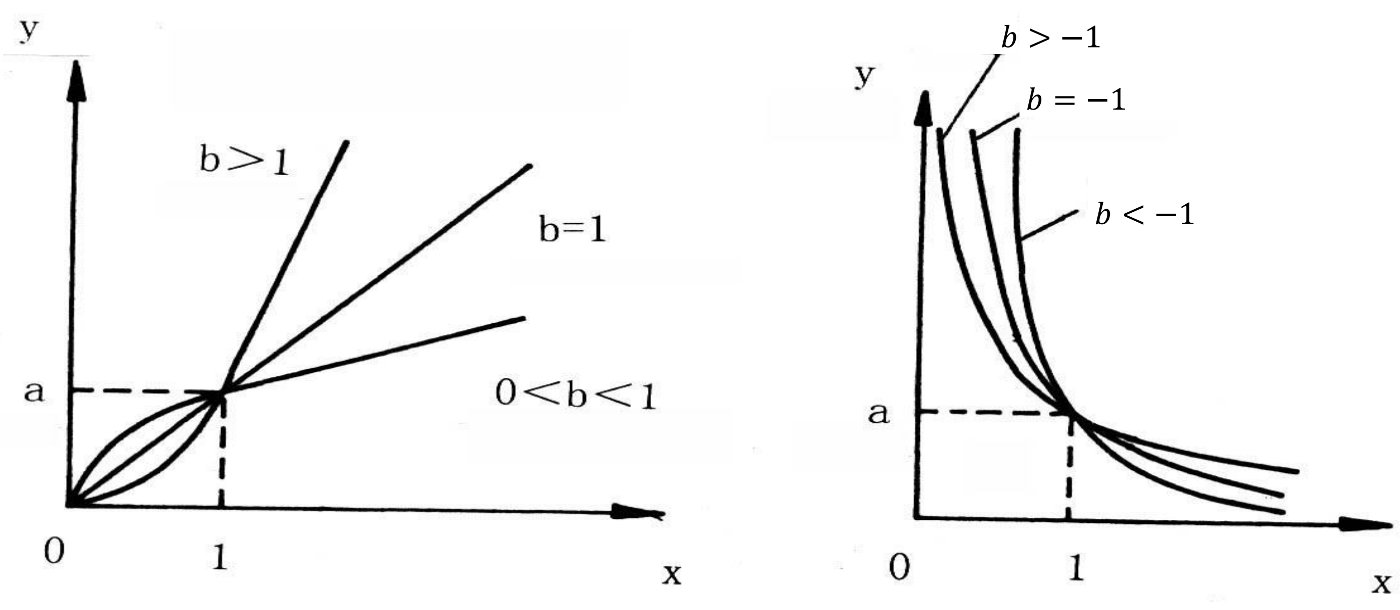
（3） 幂函数型

$$y = ax^b \qquad x > 0$$

$u = \ln y$，$v = \ln x$，得到：

$$u = \ln a + bv$$

$$y = ax^b, 令 u = \ln y, c = \ln a, v = \ln x \Rightarrow u = c + bv$$

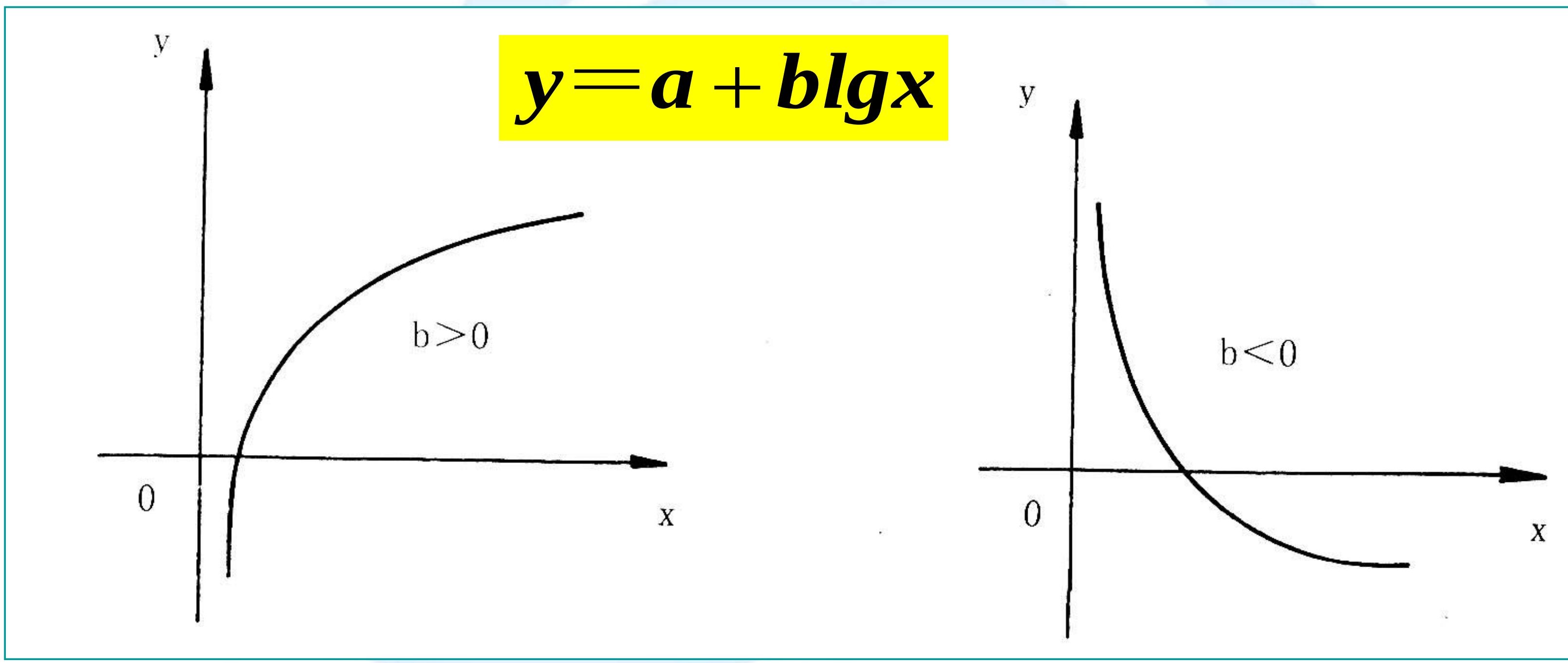


(4) 对数曲线型

$$y = a + b\lg x$$

令$u = lgx$，得到：

$$y = a + bu$$



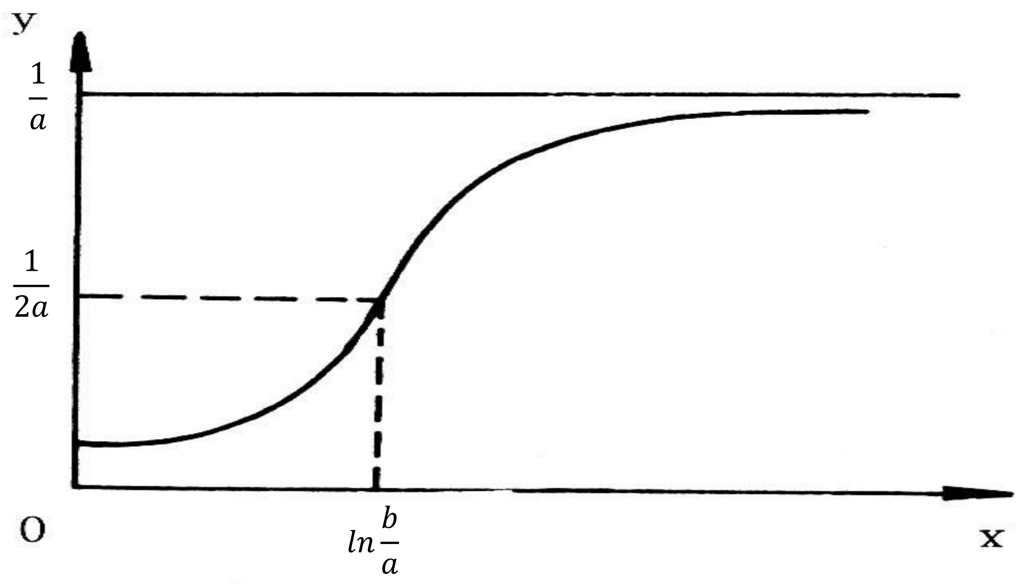
（5） S曲线型

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

令： $u=e^{-x}$ $v=1/y$

得到： $v=a+bu$

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



【例1】出钢时所用盛钢水的钢包，在使用过程中由于钢水及炉渣对耐火材料的侵蚀，容积不断增大，试验数据如表5-3所示，试找出增大容积与使用次数之间的关系。

表 5－3　钢包容积与使用次数观测数据表

使用次数 x_i	增大容积 y_i	使用次数 x_i	增大容积 y_i	使用次数 x_i	增大容积 y_i
2	6.42	7	10.00	12	10.60
3	8.20	8	9.93	13	10.80
4	9.58	9	9.99	14	10.60
5	9.50	10	10.49	15	10.90
6	9.70	11	10.59	16	10.76

$$l_{uv}=\sum_{i=1}^{n}(u_i-\overline{u})(v_i-\overline{v})=\sum_{i=1}^{n}u_iv_i-\frac{1}{n}\left(\sum_{i=1}^{n}u_i\right)\left(\sum_{i=1}^{n}v_i\right)$$

思路：先用散点图看更接近什么曲线，然后用相应方法处理。

- 根据散点图,选择双曲线

$$\frac{1}{y}=a+\frac{b}{x}$$

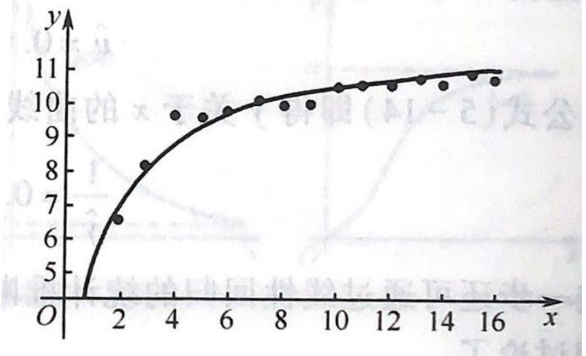


图 5－5

$$令u=\frac{1}{y},v=\frac{1}{x}\Rightarrow u=a+bv.$$

1)取倒数后得到新的数据表

$$\overline{u}=0.1031,\overline{v}=0.1587,l_{vv}=0.2065,l_{uu}=0.0038,l_{vu}=0.0271,$$

$$\hat{b}=\frac{l_{vu}}{l_{vv}}=0.1314,\hat{a}=\overline{u}-\hat{b}\overline{v}=0.0823.$$

$$\hat{u}=0.0823+0.1314v$$

$$l_{uu}=\sum_{i=1}^{n}(u_i-\overline{u})(u_i-\overline{u})=\sum_{i=1}^{n}u_i^2-\frac{1}{n}(\sum_{i=1}^{n}u_i)^2$$

$$r_1 = \frac{|l_{vu}|}{\sqrt{l_{vv}\ l_{uu}}} = 0.9686, \frac{1}{\hat{y}} = 0.0823 + \frac{0.1314}{x},$$

$$或\hat{y} = \frac{x}{0.0823x + 0.1314}.$$

以上曲线并不一定是最佳的拟合曲线,

2)根据散点图,选择倒指数曲线

$$y = ae^{\frac{b}{x}}, 令u = \ln y, v = \frac{1}{x}, c = \ln a \Rightarrow u = a + bv.$$

取对数后得到新的数据表

$$\bar{u}=2.2815,\bar{v}=0.1587,l_{vv}=0.2065,l_{uu}=0.2656,l_{uv}=-0.2293,$$

$$\hat{b}=\frac{l_{uv}}{l_{vv}}=-1.1107,\hat{c}=\bar{u}-\hat{b}\bar{v}=2.4578.$$

$$\hat{u}=2.4587-1.1107v$$

$$r_2=\frac{|l_{vu}|}{\sqrt{l_{uu}\,l_{vv}}}=0.9792,\hat{a}=e^{\hat{c}}=11.6791,$$

$$\hat{y}=11.6971e^{-\frac{1.1107}{x}}.$$

五（15 分）同一生产面积上单位产品的成本与产量间近似满足双曲线型关系 $y = a + \frac{b}{x}$，试利用下列资料求出 y 对 x 的回归曲线方程.

x_i	5.67	4.45	3.84	3.84	3.73	2.18
y_i	17.7	18.5	18.9	18.8	18.3	19.1

$t = \frac{1}{x}$

$$l_{tt} = \sum_{i=1}^{n}(t_i - \bar{t})(t_i - \bar{t}) = \sum_{i=1}^{n} {t_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} t_i\right)^2$$

$$l_{yt} = \sum_{i=1}^{n}(t_i - \bar{t})(y_i - \bar{y}) = \sum_{i=1}^{n} y_i t_i - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)\left(\sum_{i=1}^{n} t_i\right)$$

Thank you!