第9章：非线性回归

马学俊(主讲) 杜悦(助教)

苏州大学
数学科学学院

https://xuejunma.github.io/

Outline

1. 可化为线性回归的曲线回归
2. 多项式回归
3. 非线性模型

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
- $$y = \beta_0 + \beta_1 e^{bx} + \epsilon \quad (\text{b已知}) \tag{9.1}$$
  只须令$x' = e^{bx}$即可化为y对x'是线性的形式
  $$y = \beta_0 + \beta_1 x'$$
- 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。（如果b未知时，则不能通过变量替换转化为线性形式）
- $$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_p x^p + \epsilon$$
  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
  $$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_p x_p + \epsilon \tag{9.2}$$
  (9.2)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。
- 线性回归模型的“线性”是针对未知参数$\beta_i (i = 1, 2, \cdots, p)$而言的。对于回归解释变量的线性是非本质的,因为解释变量是非线性时,总可以通过变量的替换把它转化成线性的。

可化为线性回归的曲线回归

- $$y = ae^{bx}e^{\epsilon} \tag{9.3}$$
  对等式两边同时取自然对数，得：
  $$\ln y = \ln a + bx\epsilon$$
  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
  $$y' = \beta_0 + \beta_1 x + \epsilon$$
- 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
  $$y = ae^{bx} + \epsilon \tag{9.4}$$
  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。
- (9.3)式的误差项称为乘性误差项，(9.4)式的误差项称为**加性**误差项。因而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。

误差项

- 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（9.3）式$y = ae^{bx}e^{\epsilon}$简写为$y = ae^{bx}$
- (9.3)式与(9.4)式的回归参数的估计值是有差异的。
- 对误差项的形式，
  - 首先应该由数据的经济意义来确定，
  - 然后由回归拟合效果做检验。
- 过去，由于没有非线性回归软件，人们总是希望非线性回归模型可以线性化，因而误差项的形式就假定为可以把模型线性化的形式。
- 现在利用计算机软件可以容易的解决非线性回归问题，因而对误差项形式应该做正确的选择。

可化为线性回归的曲线回归

Table: 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	二次函数	$y = b_0 + b_1 t + b_2 t^2$
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形函数	$y = \exp(b_0 + b_1/t)$
Logistic	逻辑函数	$y = \frac{1}{\frac{1}{u} + b_0 b_1^t}$ u是预先给定的
Growth	增长曲线	$y = \exp(b_0 + b_1 t)$
Exponent	指数函数	$y = b_0 \exp(b_1 t)$

对以上曲线回归函数中，复合函数，增长曲线，指数函数这三个方程实际上是等价的，只是表达形式不同。

可化为线性回归的曲线回归

除了以上几种曲线回归外，另外几种其他常用的曲线回归,例如

1.双曲线函数

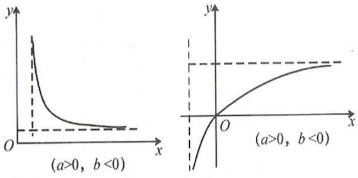
$$y = \frac{x}{ax+b}$$

或等价的表示为

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

2.S形曲线

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



(a) 双曲函数

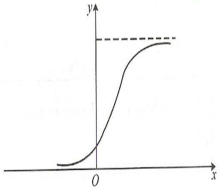


图9.1(b) S形曲线

几种常见的多项式回归模型

- 一元二次多项式模型

$$y_i = \beta_0 + \beta_1 x_i + \beta_{11} x_i^2 + \epsilon_i$$

的回归函数$y_i = \beta_0 + \beta_1 x_i + \beta_{11} x_i^2$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。

- 回归模型

$$y_i = \beta_0 + \beta_1 x_i + \beta_{111} x_i^3 + \epsilon_i$$

称为一元三次多项式模型。

几种常见的多项式回归模型

称回归模型

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

为二元二阶多项式回归模型。

- 自变量的线性项系数β_1, β_2
- 二次型系数β_{11}, β_{22}
- 交叉乘积项表示x_1与x_2的交互作用，系数β_{12}通常称为交互影响系数

非线性最小二乘

非线性回归模型一般可记为

$$y_i = f(\boldsymbol{x}_i, \boldsymbol{\theta}) + \epsilon_i, \quad i = 1, 2, \cdots, n \tag{9.8}$$

其中，y_i是因变量，非随机向量$\boldsymbol{x}_i = (x_{i1}, x_{i2}, \cdots, x_{ik})'$是自变量，$\boldsymbol{\theta} = (\theta_0, \theta_1, \cdots, \theta_p)'$是未知参数向量，$\epsilon_i$是随机误差项并且满足独立同分布假定，即

$$\begin{cases} E(\epsilon_i) = 0, \quad i = 1, 2, \ldots, n \\ cov(\epsilon_i, \epsilon_j) = \begin{cases} \sigma^2, i = j \\ 0, i \neq j \end{cases} \quad i, j = 1, 2, \ldots, n \end{cases}$$

如果$f(\boldsymbol{x}_i, \boldsymbol{\theta}) = \theta_0 + x_{i1}\theta_1 + x_{i2}\theta_2 + \cdots + x_{ip}\theta_p$，那么式(9.8)就是前面讨论的线性模型，而且必然有$k = p$；对于一般情况的非线性模型，参数的数目与自变量的数目并没有一定的对应关系，不要求$k = p$。

非线性最小二乘

对非线性回归模型式(9.8)，仍使用最小二乘法估计参数$\boldsymbol{\theta}$，即求使

$$Q(\boldsymbol{\theta}) = \sum_{i=1}^{n}(y_i - f(\boldsymbol{x}_i, \boldsymbol{\theta}))^2$$

达到最小的$\hat{\boldsymbol{\theta}}$，称$\hat{\boldsymbol{\theta}}$为线性最小二乘估计。在假定$f$函数对参数$\boldsymbol{\theta}$连续可微时，可以利用微分法建立正规方程组，求使$Q(\boldsymbol{\theta})$达到最小的$\hat{\boldsymbol{\theta}}$。将$Q$函数对参数$\theta_j$求偏导，并令其为0，得$p+1$个方程

$$\frac{\partial Q}{\partial \theta_j}\Big|_{\theta_j=\hat{\theta}_j} = -2\sum_{i=1}^{n}(\boldsymbol{y}_i - f(\boldsymbol{x}_i, \hat{\boldsymbol{\theta}}))\frac{\partial f}{\partial \theta_j}\Big|_{\theta_j=\hat{\theta}_j} = 0, j = 1, 2, \cdots, p \quad (9.10)$$

非线性最小二乘估计$\hat{\boldsymbol{\theta}}$就是式(9.10)得解，式(9.10)称为非线性最小二乘估计得正规方程组，它是未知参数得非线性方程组。一般用Newton迭代法求解此正规方程组，也可以直接极小化残差平方和$Q(\boldsymbol{\theta})$，求出未知参数$\boldsymbol{\theta}$的非线性最小二乘估计值$\hat{\boldsymbol{\theta}}$

非线性最小二乘

- 对于非线性最小二乘估计，我们仍然需要做参数的区间估计、显著性检验、回归方程的显著性检验等回归诊断，这需要知道有关统计量的分布。
- 在非线性最小二乘中，一些精确分布是很难得到的，在大样本时，可以得到近似的分布
- 计算机软件在求出参数$\boldsymbol{\theta}$的非线性最小二乘估计值的同时，还给出近似的参数的区间估计、显著性检验、回归方程的显著性检验等回归诊断。
- 在非线性回归中，平方和分解式$SST = SSR + SSE$不再成立。类似于线性回归中的复判定系数，定义非线性回归的相关比为：

$$R^2 = 1 - \frac{SSE}{SST}$$

相关比也称为相关指数。

作业

- p.228 9.5