应用数理统计

Ch5 回归分析

——————多元线性回归

2014年6月25日

多元线性回归的数学模型

设自变量$x_1, x_2, \cdots, x_m$与因变量y之间，有下列关系

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \varepsilon$$

其中，$\beta_0, \beta_1, \cdots, \beta_m$是常数，$\varepsilon \sim N(0, \sigma^2)$是表示误差的随机变量。

对$x_1, x_2, \cdots, x_m, y$进行n次观测，得到一组观测值：

$$(x_{i1}, x_{i2}, \cdots, x_{im}, y_i),\quad i = 1, 2, \cdots, n$$

$$\begin{cases} y_1 = \beta_0 + \beta_1 x_{11} + \cdots + \beta_m x_{1m} + \varepsilon_i \\ \qquad\qquad \vdots \\ y_n = \beta_0 + \beta_1 x_{n1} + \cdots + \beta_m x_{nm} + \varepsilon_n \end{cases}$$

$\varepsilon_i \sim N(0, \sigma^2)$ ，$i = 1, 2, \cdots, n$ ，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$相互独立

为了简单起见，我们将它写成矩阵向量形式。

令$Y=\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$，$X=\begin{bmatrix} 1 & x_{11} & \cdots & x_{1m} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nm} \end{bmatrix}$，$\beta=\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$，$e=\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$

$$\begin{cases} y_1=\beta_0+\beta_1 x_{11}+\cdots+\beta_m x_{1m}+\varepsilon_1 \\ \qquad\qquad\vdots \\ y_n=\beta_0+\beta_1 x_{n1}+\cdots+\beta_m x_{nm}+\varepsilon_n \end{cases} \Longrightarrow Y=X\beta+e,\quad e\sim N_n(\mathbf{0},\sigma^2 I)$$

$N_n(\mathbf{0},\sigma^2 I)$表示$n$元正态分布，$\mathbf{0}=\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$是数学期望，$\sigma^2 I=\begin{bmatrix} \sigma^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2 \end{bmatrix}$是协方差矩阵

问题： 已知数据矩阵$X=\begin{bmatrix}1 & x_{11} & \cdots & x_{1m}\\ \vdots & \vdots & & \vdots\\ 1 & x_{n1} & \cdots & x_{nm}\end{bmatrix}$和数据向量$\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}$,

求β的最小二乘估计$\hat{\beta}=\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \vdots\\ \hat{\beta}_m\end{bmatrix}$，使得下列平方和达到最小：

$$Q=\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_{i1}-\cdots-\beta_m x_{im})^2$$

分析推导：

Q是$\beta_0,\beta_1,\cdots,\beta_m$的函数，所以这是一个多元函数求最小极值的问题

$$\left\{\begin{array}{l} \dfrac{\partial Q}{\partial \beta_0} = 0 \\ \dfrac{\partial Q}{\partial \beta_1} = 0 \\ \quad \cdots\cdots \\ \dfrac{\partial Q}{\partial \beta_m} = 0 \end{array}\right. \Rightarrow \quad \hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_m \end{bmatrix} \Rightarrow \quad \hat{\beta} = (\mathrm{X}^T \mathrm{X})^{-1} X^T Y$$

$$Q = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_m x_{im})^2$$

$$= \begin{bmatrix} y_1 - \beta_0 - \beta_1 x_{11} - \cdots - \beta_m x_{1m} \\ \vdots \\ y_n - \beta_0 - \beta_1 x_{n1} - \cdots - \beta_m x_{nm} \end{bmatrix}^T \begin{bmatrix} y_1 - \beta_0 - \beta_1 x_{11} - \cdots - \beta_m x_{1m} \\ \vdots \\ y_n - \beta_0 - \beta_1 x_{n1} - \cdots - \beta_m x_{nm} \end{bmatrix}$$

$$= (Y - X\beta)^T (Y - X\beta) = Y^T Y - 2\beta^T X^T Y + \beta^T X^T X \beta$$

$$\frac{\partial Q}{\partial \beta}=\frac{\partial}{\partial \beta}(Y^TY-2\beta^TX^TY+\beta^TX^TX\beta)=\mathbf{0}-2X^TY+2X^TX\beta=\mathbf{0}$$

用到: $\frac{\partial}{\partial x}A=\mathbf{0},\qquad \frac{\partial}{\partial x}(x^TA)=A,\qquad \frac{\partial}{\partial x}(x^TAx)=(A+A^T)x$

$$\frac{\partial Q}{\partial \beta}=\frac{\partial}{\partial \beta}(Y^TY-2\beta^TX^TY+\beta^TX^TX\beta)=\mathbf{0}-2X^TY+2X^TX\beta=\mathbf{0}$$

$X^TX\beta=X^TY$ -------- 正规方程

$\hat{\beta}=(\mathrm{X}^T\mathrm{X})^{-1}X^TY$ -------- $\hat{\beta}=\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \vdots\\ \hat{\beta}_m\end{bmatrix}$的最小二乘估计

$$\left.\begin{array}{l} X^T X \beta = X^T Y \\ \hat{Y} = \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{bmatrix} = X \hat{\beta} \end{array}\right\} \Rightarrow \boxed{X^T \hat{Y} = X^T X \hat{\beta} = X^T Y}$$

等式两边的第一个分量相等

$$\boxed{\sum \hat{y}_i = \sum y_i}$$

$$\Downarrow$$

$$\boxed{\hat{y}_i\text{的均值} \;=\; y_i\text{的均值 } \bar{y}}$$

$$SS_e = Q_{\min} = \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \cdots - \hat{\beta}_m x_{im})^2$$

$$= Y^T Y - 2\hat{\beta}^T X^T Y + \hat{\beta}^T X^T X \hat{\beta}$$

由于$\hat{\beta} = (X^T X)^{-1} X^T Y$，$\hat{\beta}^T X^T Y = \hat{\beta}^T X^T X \hat{\beta}$，因此

$$SS_e = Y^T Y - \hat{\beta}^T X^T Y \quad 或 \quad SS_e = Y^T Y - \hat{\beta}^T X^T X \hat{\beta}$$

$$\hat{\sigma} = \sqrt{\frac{SS_e}{n-m-1}}$$

为估计标准差（或残差标准差）

$$r = \sqrt{1 - \frac{SS_e}{L_{yy}}} = ? = \sqrt{\frac{SS_R}{SS_T}}$$

为多重相关系数

L_{yy}

如果记： $SS_T = \sum_{i=1}^{n}(y_i - \bar{y})^2$ ———总离差平方和

$$SS_R = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \text{———回归平方和}$$

及前面讲到的：

$$SS_e = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \text{———残差平方和}$$

则可以证明：

$$SS_T = SS_R + SS_e \text{———离差分解公式}$$

$$r=\sqrt{1-\frac{SS_e}{L_{yy}}}=\sqrt{\frac{L_{yy}-SS_e}{L_{yy}}}=\sqrt{\frac{SS_T-SS_e}{SS_T}}=\sqrt{\frac{SS_R}{SS_T}}$$

$r^2=\dfrac{SS_R}{SS_T}$ ---判定系数, 反映回归方程对观测数据的拟和程度

复相关系数 $r=\sqrt{\dfrac{SS_R}{SS_T}}$ 是因变元y与一组自变元 $x_1, x_2, ..., x_m$的相关系数,这如何理解呢?

事实上, $r=\sqrt{\dfrac{SS_R}{SS_T}}$ 就是$Y=\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ 与其回归估计值$\hat{Y}=\begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{bmatrix}$的简单相关系数

证明: $Y=\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ 与 $\hat{Y}=\begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{bmatrix}$的简单相关系数为:

$$\frac{\sum (y_i - \overline{y})(\hat{y}_i - \overline{y})}{\sqrt{\sum (y_i - \overline{y})^2 \sum (\hat{y}_i - \overline{y})^2}} = \frac{\sum (y_i - \overline{y})(\hat{y}_i - \overline{y})}{\sqrt{SS_T SS_R}}$$

要证: $\frac{\sum (y_i - \overline{y})(\hat{y}_i - \overline{y})}{\sqrt{SS_T SS_R}} = \sqrt{\frac{SS_R}{SS_T}}$

$$\Leftrightarrow \sum (y_i - \overline{y})(\hat{y}_i - \overline{y}) = SS_R$$

$$\Leftrightarrow \sum [(y_i - \overline{y})-(\hat{y}_i - \overline{y})](\hat{y}_i - \overline{y}) = 0$$

$$\Leftrightarrow \sum y_i \hat{y}_i - \sum \hat{y}_i^2 = 0$$

$$\Leftrightarrow Y^T \hat{Y} = \hat{Y}^T \hat{Y} \Leftrightarrow Y^T (X\hat{\beta}) = (X\hat{\beta})^T (X\hat{\beta})$$

$$\boxed{\hat{\beta} = (X^T X)^{-1} X^T Y}$$

例 1 设有多元线性回归模型

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i \text{ ，} i = 1, 2, 3, 4 \text{ ，}$$

其中，$\beta_0, \beta_1, \beta_2$ 是未知常数，$\varepsilon_i \sim N(0, \sigma^2)$，$i = 1, 2, 3, 4$ ，相互独立。并且已知

$$x_{11} = 0 \text{ ，} x_{12} = 1 \text{ ，} y_1 = -2 \text{ ，}$$

$$x_{21} = 1 \text{ ，} x_{22} = -1 \text{ ，} y_2 = 3 \text{ ，}$$

$$x_{31} = 0 \text{ ，} x_{32} = -1 \text{ ，} y_3 = 2 \text{ ，}$$

$$x_{41} = -1 \text{ ，} x_{42} = 1 \text{ ，} y_4 = 1 \text{ 。}$$

求 $\beta_0, \beta_1, \beta_2$ 的最小二乘估计 $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ 。

解： $$X=\begin{bmatrix}1 & x_{11} & x_{12}\\ 1 & x_{21} & x_{22}\\ 1 & x_{31} & x_{32}\\ 1 & x_{41} & x_{42}\end{bmatrix}=\begin{bmatrix}1 & 0 & 1\\ 1 & 1 & -1\\ 1 & 0 & -1\\ 1 & -1 & 1\end{bmatrix},\quad Y=\begin{bmatrix}y_1\\ y_2\\ y_3\\ y_4\end{bmatrix}=\begin{bmatrix}-2\\ 3\\ 2\\ 1\end{bmatrix}$$

$$\begin{bmatrix}\hat\beta_0\\ \hat\beta_1\\ \hat\beta_2\end{bmatrix}=\hat\beta=(X^TX)^{-1}X^TY=\begin{bmatrix}1/4 & 0 & 0\\ 0 & 1 & 1/2\\ 0 & 1/2 & 1/2\end{bmatrix}\begin{bmatrix}4\\ 2\\ -6\end{bmatrix}=\begin{bmatrix}1\\ -1\\ -2\end{bmatrix}$$

所以$\beta_0, \beta_1, \beta_2$的最小二乘估计为$\beta_0=1, \beta_1=-1, \beta_2=-2$

多元线性回归中统计量的分布

$$Y = X\beta + e, e \sim N_n(0, \sigma^2 I) \quad Y = X\beta + e, e \sim N_n(X\beta, \sigma^2 I)$$

$$\hat{\beta} \sim N_{m+1}(\beta, \sigma^2 (X^T X)^{-1})$$

$$\frac{SSe}{\sigma^2} \sim \chi^2 (n - m - 1)$$

$$\hat{\beta}_j \sim N(\beta_j, c_{jj}\sigma^2), \quad (j=1,2,...,m)$$

其中 $(X^T X)^{-1} = \left(c_{ij}\right)_{(m+1)\times(m+1)}$

$$\frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}\sqrt{c_{jj}}} \sim t(n-m-1)$$

定理8 $E(\hat{\beta})=\beta$，$D(\hat{\beta})=\sigma^2(X^TX)^{-1}$

证： 由于$Y\sim N_n(X\beta,\sigma^2 I)$,所以$E(Y)=X\beta$，$D(Y)=\sigma^2 I$，因此有

$$E(\hat{\beta})=E[(X^TX)^{-1}X^TY]=(X^TX)^{-1}X^TE(Y)=(X^TX)^{-1}X^TX\beta=\beta$$

$$D(\hat{\beta})=D[(X^TX)^{-1}X^TY]=(X^TX)^{-1}X^TD(Y)X(X^TX)^{-1}$$

$$=(X^TX)^{-1}X^T\sigma^2 I\ X(X^TX)^{-1}$$

$$=\sigma^2(X^TX)^{-1}X^TX(X^TX)^{-1}=\sigma^2(X^TX)^{-1}$$

定理9 $\hat{\beta} \sim N_m(\beta, \sigma^2 (X^T X)^{-1})$

证： 因为Y服从正态分布，

而$\hat{\beta} = (X^T X)^{-1} X^T Y$是$Y$的线性函数，

所以$\hat{\beta}$也服从正态分布，

由定理8可知$E(\hat{\beta}) = \beta$，$D(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$

因此，$\hat{\beta} \sim N_m(\beta, \sigma^2 (X^T X)^{-1})$

定理10 $\frac{SS_e}{\sigma^2} \sim \chi^2(n-m-1)$，而且$SS_e$与$\hat{\beta}$相互独立

证： 因为$Y = X\beta + e$，所以$e = Y - X\beta$

$$e^T e = (Y - X\beta)^T (Y - X\beta) = Y^T Y - 2\beta^T X^T Y + \beta^T X^T X\beta$$

$$= Y^T Y - \hat{\beta}^T X^T X\hat{\beta} + \hat{\beta}^T X^T X\hat{\beta} - 2\beta^T X^T X\hat{\beta} + \beta^T X^T X\beta$$

$$= SS_e + (\hat{\beta} - \beta)^T X^T X(\hat{\beta} - \beta)$$

因为$\varepsilon_i \sim N(0, \sigma^2)$，$i = 1, 2, \cdots, n$相互独立，

所以$\frac{\varepsilon_i}{\sigma} \sim N(0, \sigma^2)$，$i = 1, 2, \cdots, n$相互独立

$$\sum_{i=1}^{n} \left(\frac{\varepsilon_i}{\sigma}\right)^2 = \frac{\sum_{i=1}^{n} \varepsilon_i^2}{\sigma^2} = \frac{e^T e}{\sigma^2} = \frac{SS_e + (\hat{\beta} - \beta)^T X^T X(\hat{\beta} - \beta)}{\sigma^2}$$

$$=\frac{SS_e}{\sigma^2}+\frac{(\hat{\beta}-\beta)^T X^T X(\hat{\beta}-\beta)}{\sigma^2}=Q_1+Q_2$$

其中$Q_1=\frac{SS_e}{\sigma^2}=\frac{\sum_{i=1}^{n}(y_i-\hat{\beta}_0-\hat{\beta}_1 x_{i1}-\cdots-\hat{\beta}_m x_{im})^2}{\sigma^2}=\frac{(Y-X\hat{\beta})^T(Y-X\hat{\beta})}{\sigma^2}$

是n项的平方和，但这n项又满足$m+1$个线性关系式

$$X^T(Y-X\hat{\beta})=X^TY-X^TX\hat{\beta}=X^TY-X^TX(X^TX)^{-1}X^TY=\mathbf{0}$$

所以Q_1的自由度$f_1=n-m-1$

$$Q_2=\frac{(\hat{\beta}-\beta)^T X^T X(\hat{\beta}-\beta)}{\sigma^2}\text{是}m+1\text{项平方和}$$

因为X^TX的秩是$m+1$, 故存在$m+1$阶方阵方阵P使得:

$(\hat{\beta}-\beta)^T X^T X(\hat{\beta}-\beta)=(\hat{\beta}-\beta)^T P^T P(\hat{\beta}-\beta)$,所以, Q_2的自由度 $f_2=m+1$

因为$f_1 + f_2 = (n-m-1)+(m+1) = n$,

所以由定理2.7（Cochran定理）可知：

$$Q_1 = \frac{SS_e}{\sigma^2} \sim \chi^2(n-m-1),$$

$$Q_2 = \frac{(\hat{\beta}-\beta)^T X^T X(\hat{\beta}-\beta)}{\sigma^2} \sim \chi^2(m+1)$$

而且Q_1，Q_2相互独立，所以SS_e与$\hat{\beta}$ 相互独立

定理11 $E(SS_e)=(n-m-1)\sigma^2$，$E(\hat{\sigma}^2)=\sigma^2$

证： 若$\xi \sim \chi^2(n)$，则$E\xi=n$

由定理10可知，$\dfrac{SS_e}{\sigma^2}\sim\chi^2(n-m-1)$，所以$E(\dfrac{SS_e}{\sigma^2})=n-m-1$ ，因此

$$E(SS_e)=E(\frac{SS_e}{\sigma^2}\sigma^2)=E(\frac{SS_e}{\sigma^2})\sigma^2=(n-m-1)\sigma^2$$

$$E(\hat{\sigma}^2)=E(\frac{SS_e}{n-m-1})=\frac{E(SS_e)}{n-m-1}=\frac{(n-m-1)\sigma^2}{n-m-1}=\sigma^2$$

即$\hat{\sigma}^2=\dfrac{SS_e}{n-m-1}$是$\sigma^2$的无偏估计

定理12 $\dfrac{\hat{\beta}_j-\beta_j}{\hat{\sigma}\sqrt{c_{jj}}}\sim t(n-m-1)$，$j=1,2,\cdots,m$

其中C_{jj}是矩阵$(X^TX)^{-1}=\begin{bmatrix} c_{00} & \cdots & \cdots & c_{0m} \\ \vdots & c_{11} & & \vdots \\ \vdots & & \ddots & \vdots \\ c_{m0} & \cdots & \cdots & c_{mm} \end{bmatrix}$的第$j+1$个对角元素

证： 由定理9可知$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_m \end{bmatrix}=\hat{\beta}\sim N_m(\beta,\sigma^2(X^TX)^{-1})$，其中

数学期望$\beta=\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$，协方差$\sigma^2(X^TX)^{-1}=\sigma^2\begin{bmatrix} c_{00} & \cdots & \cdots & c_{0m} \\ \vdots & c_{11} & & \vdots \\ \vdots & & \ddots & \vdots \\ c_{m0} & \cdots & \cdots & c_{mm} \end{bmatrix}$

所以$\hat{\beta}_j\sim N(\beta_j,\sigma^2 c_{jj})$，即$\dfrac{\hat{\beta}_j-\beta_j}{\sigma\sqrt{c_{jj}}}\sim N(0,1)$

$$\frac{SS_e}{\sigma^2} \sim \chi^2(n-m-1)$$，而且$\hat{\beta}$与SS_e相互独立，

$$\frac{\hat{\beta}_j-\beta_j}{\sigma\sqrt{c_{jj}}}=\frac{\dfrac{\hat{\beta}_1-\beta_1}{\sigma\sqrt{c_{jj}}}}{\sqrt{\dfrac{SS_e}{\sigma^2}\Big/(n-m-1)}} \sim t(n-m-1)$$

多元线性回归中的假设检验

（1）检验全部自变量$x_1, x_2, \cdots, x_m$与因变量y之间是否有统计线性相关关系

相当于检验 回归模型 $y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \varepsilon$， $\varepsilon \sim N(0, \sigma^2)$
中自变元的系数全部为零, 即 $H_0 : \beta_1 = \beta_2 = \cdots = \beta_m = 0$

因为$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \varepsilon$， $\varepsilon \sim N(0, \sigma^2)$

如果假设H_0为真， $\beta_1 = \beta_2 = \cdots = \beta_m = 0$ ， 则有$y = \beta_0 + \varepsilon$
说明$x_1, x_2, \cdots, x_m$与y无关， 反之说明$x_1, x_2, \cdots, x_m$与y有关

问题 已知有一组观测值 $(x_{i1}, x_{i2}, \cdots, x_{im}, y_i)$，$i = 1, 2, \cdots, n$，设

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_m x_{im} + \varepsilon_i \text{ ， } \varepsilon_i \sim N(0, \sigma^2) \text{ ， } i = 1, 2, \cdots, n$$

$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 相互独立，要检验 H_0：$\beta_1 = \beta_2 = \cdots = \beta_m = 0$。

在 H_0 为真时，SS_R 与 SS_e 相互独立，$\dfrac{SS_R}{\sigma^2} \sim \chi^2(m)$,

$$F = \frac{SS_R / m}{SS_e / (n-m-1)} \sim F(m, n-m-1).$$

若 H_0 不真，则 F 的值会偏大

当 $F > F_{1-\alpha}(m, n-m-1)$ 时拒绝 H_0，否则接受 H_0

(2) 检验某一个自变量 x_j 是否与因变量 y 统计线性相关

因为，在多元线性回归方程中，自变量不止一个，即使对全体自变量来说，已经可以肯定它们与因变量 y 有关，但是，对其中某一个自变量 x_j 来说，却不能保证它一定与 y 有关，所以，还需要作这方面的检验。

不难看出，检验 x_j 与 y 之间是否统计线性相关，相当于要检验这样一个假设 H_{0j}： $\beta_j = 0$ 。

$\beta_j = 0$ 时，$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{j-1} x_{i\,j-1} + 0 + \beta_{j+1} x_{i\,j+1} + \cdots + \beta_m x_{im} + \varepsilon_i$ ，y 与 x_j 无关。

问题 已知有一组观测值 $(x_{i1}, x_{i2}, \cdots, x_{im}, y_i)$ ，$i = 1, 2, \cdots, n$ ，设

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_m x_{im} + \varepsilon_i \text{ ，} \varepsilon_i \sim N(0, \sigma^2) \text{ ，} i = 1, 2, \cdots, n$$

$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 相互独立，要检验 H_{0j}： $\beta_j = 0$ 。

检验方法一 (t 检验)

$$\frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}\sqrt{c_{jj}}} \sim t(n-m-1)$$

若H_{0j}为真$\beta_j = 0$，有$T_j = \dfrac{\hat{\beta}_j}{\hat{\sigma}\sqrt{c_{jj}}} = \dfrac{\hat{\beta}_j - \beta_j}{\hat{\sigma}\sqrt{c_{jj}}} \sim t(n-m-1)$

当$\left|T_j\right| > t_{1-\alpha/2}(n-m-1)$时拒绝$H_{0j}$，否则接受$H_{0j}$

检验方法二 (F 检验)

$$F_j = T_j^2 = \left(\frac{\hat{\beta}_j}{\hat{\sigma}\sqrt{c_{jj}}}\right)^2 = \frac{\hat{\beta}_j^2 \big/ c_{jj}}{SS_e \big/ (n-m-1)}$$

若H_{0j}为真$\beta_j = 0$, 有$F_j = T_j^2 \sim \mathrm{F}(1, n-m-1)$

若$H_{0j}: \beta_j=0$不真，则T_j的绝对值会偏大，则$F_j = T_j^2$ 的值也会偏大

当$F > F_{1-\alpha}(m, n-m-1)$时拒绝$H_{0j}$

例2 对某种化工产品的产量 y（单位：kg），生产时的处理压力 x_1（单位：大气压）和温度 x_2 作测量，得数据如下：

压力 x_{i1}	温度 x_{i2}	产量 y_i	压力 x_{i1}	温度 x_{i2}	产量 y_i	压力 x_{i1}	温度 x_{i2}	产量 y_i
6.8	665	40	8.4	670	57	9.5	685	59
7.2	685	49	8.6	675	58	9.7	700	67
7.6	690	55	8.8	690	62	10.0	650	56
8.0	700	63	9.1	700	65	10.3	690	72
8.2	695	65	9.3	680	58	10.5	670	68

设有 $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$，$\varepsilon_i \sim N(0, \sigma^2)$，$i = 1, 2, \cdots, 15$，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{15}$ 相互独立。

求：（1）$\beta_0, \beta_1, \beta_2$ 的最小二乘估计 $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$；

（2）残差平方和 SS_e，估计的标准差 $\hat{\sigma}$，多重相关系数 r；

（3）检验 H_0：$\beta_1 = \beta_2 = 0$（显著水平 $\alpha = 0.05$）；

（4）分别检验 H_{01}：$\beta_1 = 0$ 和 H_{02}：$\beta_2 = 0$（显著水平 $\alpha = 0.05$）

解： 利用可作多元线性回归的计算机软件，求得：

（1）$\hat{\beta}_0=-200.455$，$\hat{\beta}_1=5.68337$，$\hat{\beta}_2=0.307528$，所以，回归方程为

$$\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x_1+\hat{\beta}_2 x_2=-200.455+5.68337x_1+0.307528x_2$$

（2）残差平方和$SS_e=123.806$，估计的标准差$\hat{\sigma}=3.212$，

多重相关系数$r=0.9285$

（3）检验H_0：$\beta_1=\beta_2=0$的统计量$F=37.50$ ，

对$\alpha=0.05$，查F分布表，可得分为数：

$F_{1-\alpha}(m,n-m-1)=F_{0.95}(2,12)=3.89$ ，

因为 $F=37.50>3.89$，所以拒绝H_0：$\beta_1=\beta_2=0$，

说明自变量x_1,x_2与因变量y之间有显著的统计线性相关关系

（4）检验H_{01}：$\beta_1 = 0$的统计量$F_1 = 54.10$

$F_{1-\alpha}(1, n-m-1) = F_{0.95}(1\ ,12) = 4.75$,

因为$F_1 = 54.10 > 4.75$，所以拒绝H_{01}：$\beta_2 = 0$

检验H_{02}：$\beta_2 = 0$的统计量$F_2 = 27.19$

$F_{1-\alpha}(1, n-m-1) = F_{0.95}(1\ ,12) = 4.75$

因为$F_2 = 27.19 > 4.75$，所以拒绝H_{02} ：$\beta_2 = 0$

线性回归的方差分析表

方差源	离差平方和	自由度	F值	p值
回归	SS_R	$f_R = m$	$F = \dfrac{SS_R / f_R}{SS_e / f_E}$	$p = P\{F(m, n-m-1) > F\}$
残差	SS_e	$f_E = n-m-1$		
总计	SS_T	$f_T = n-1$		

统计量观测值F>分位数 $\Longleftrightarrow$ p值小于给定显著性水平

例 3 设 θ_i，$i=1,2,3$ 是一个三角形的三个角，它们的观测值分别为 y_i，$i=1,2,3$，每次观测的误差 $\varepsilon_i \sim N(0,\sigma^2)$，$i=1,2,3$，$\varepsilon_1,\varepsilon_2,\varepsilon_3$ 相互独立。求：

（1）θ_i，$i=1,2,3$ 的估计值；（2）σ^2 的估计值；（3）检验 H_{0i}：$\theta_i=\theta_{i0}$ 的检验方法。

解 因为 $\theta_1+\theta_2+\theta_3=\pi$，只有 $m=2$ 个独立的未知参数，所以，只要考虑 θ_1,θ_2 的回归就可以了

回归方程为 $y=\theta_1 x_1+\theta_2 x_2+\varepsilon$ ，$n=3$ 次观测对应于 $\begin{cases} y_1=\theta_1+\varepsilon_1 \\ y_2=\theta_2+\varepsilon_2 \\ \pi-y_3=\theta_1+\theta_2+\varepsilon_3 \end{cases}$

即有 $Y=X\theta+e$ ，其中 $Y=\begin{bmatrix} y_1 \\ y_2 \\ \pi-y_3 \end{bmatrix}$，$X=\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$，$\theta=\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$，$e=\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$。

这是一个不含常数项的多元线性回归模型，可以像含有常数项的多元线性回归模型一样

唯一不同是**:**含有常数项的多元线性回归模型中**m+1,**在不含常数的多元线性回归

$$(m+1) \longrightarrow m$$

（1）

$$\begin{bmatrix}\hat{\theta}_1\\ \hat{\theta}_2\end{bmatrix}=\hat{\theta}=(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}Y=\left(\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}\right)^{-1}\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\\pi-y_3\end{bmatrix}$$

$$=\begin{bmatrix}2&1\\1&2\end{bmatrix}^{-1}\begin{bmatrix}y_1+\pi-y_3\\y_2+\pi-y_3\end{bmatrix}=\begin{bmatrix}2/3&-1/3\\-1/3&2/3\end{bmatrix}\begin{bmatrix}y_1+\pi-y_3\\y_2+\pi-y_3\end{bmatrix}=\begin{bmatrix}\dfrac{\pi+2y_1-y_2-y_3}{3}\\\dfrac{\pi-y_1+2y_2-y_3}{3}\end{bmatrix}$$

（2）残差平方和

$$SS_e=(y_1-\hat{\theta}_1)^2+(y_2-\hat{\theta}_2)^2+(y_3-\hat{\theta}_3)^2$$

$$=(y_1-y_1-\frac{\pi-y_1-y_2-y_3}{3})^2+(y_2-y_2-\frac{\pi-y_1-y_2-y_3}{3})^2+(y_3-y_3-\frac{\pi-y_1-y_2-2y_3}{3})^2$$

$$=3(-\frac{\pi-y_1-y_2-y_3}{3})^2=\frac{(\pi-y_1-y_2-y_3)^2}{3}\text{。}$$

可以证明 $\frac{SS_e}{\sigma^2}\sim\chi^2(n-m)$，所以 $E(\frac{SS_e}{\sigma^2})=n-m$，即有 $E(\frac{SS_e}{n-m})=\sigma^2$

$$\hat{\sigma}^2=\frac{SS_e}{n-m}=\frac{1}{3-2}\cdot\frac{(\pi-y_1-y_2-y_3)^2}{3}=\frac{(\pi-y_1-y_2-y_3)^2}{3}$$

（3）因为 $\begin{bmatrix}\hat{\theta}_1\\ \hat{\theta}_1\end{bmatrix} \sim N\left(\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}, \sigma^2(X^{\mathrm{T}}X)^{-1}\right) = N\left(\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}, \sigma^2\begin{bmatrix}2/3 & -1/3\\ -1/3 & 2/3\end{bmatrix}\right)$ ，所以

$$\hat{\theta}_i \sim N(\theta_i, \frac{2}{3}\sigma^2)\ ,\quad \frac{\hat{\theta}_i - \theta_i}{\sigma}\sqrt{\frac{3}{2}} \sim N(0,1)\ ,\quad i = 1, 2\ 。$$

再加上有 $\dfrac{(n-m)\hat{\sigma}^2}{\sigma^2} = \dfrac{SS_e}{\sigma^2} \sim \chi^2(n-m)$，且与上式独立，所以当 H_0： $\theta_i = \theta_{i0}$ 为真时，有

$$T_i = \frac{\hat{\theta}_i - \theta_{i0}}{\hat{\sigma}}\sqrt{\frac{3}{2}} = \frac{\hat{\theta}_i - \theta_i}{\hat{\sigma}}\sqrt{\frac{3}{2}} = \frac{\dfrac{\hat{\theta}_i - \theta_i}{\sigma}\sqrt{\dfrac{3}{2}}}{\sqrt{\dfrac{(n-m)\hat{\sigma}^2}{\sigma^2}\Big/(n-m)}} \sim t(n-m) = t(1)\ ,\quad i = 1, 2\ 。$$

因此，检验假设 H_0： $\theta_i = \theta_{i0}$ 的 t 分布检验方法为：

求出统计量 $T_i = \dfrac{\hat{\theta}_i - \theta_{i0}}{\hat{\sigma}}\sqrt{\dfrac{3}{2}} = \dfrac{y_i + \dfrac{\pi - y_1 - y_2 - y_3}{3} - \theta_{i0}}{\sqrt{\dfrac{(\pi - y_1 - y_2 - y_3)^2}{3}}}\sqrt{\dfrac{3}{2}} = \dfrac{3y_i - 3\theta_{i0} + \pi - y_1 - y_2 - y_3}{\sqrt{2}\,|\pi - y_1 - y_2 - y_3|}$ 的值，对给定的显著水平 α，查 t 分布表，可得 $t_{1-\alpha/2}(n-m) = t_{1-\alpha/2}(1)$ ， 当 $|T_i| > t_{1-\alpha/2}(1)$ 时拒绝 H_0： $\theta_i = \theta_{i0}$ ，否则接受 H_0。

思考题

1）以上多元回归的方法和结论有个假设，即X'X可逆，这等价于X满秩。如果这个条件不满足怎么办？

-----------参见5.5节　逐步回归方法

2）多元线性回归的第二个前提是误差服从正态分布，如果不是正态分布怎么办？

-------参阅基于指数分布族的广义回归

3）变元间是非线性关系，如何求回归方程？

5.4 非线性回归

可以化为线性的非线性回归——广义线性回归

例1 某零件上有一条曲线，可以近似看作是一条抛物线，为了在数控机床上加工这一零件，在曲线上测得n个点的坐标(x_i, y_i)，$i=1, 2, \cdots, n$，要求从这n个点的坐标出发，求出曲线的函数表达式。

显然，这是一个回归分析问题，由于曲线可以近似看作是一条抛物线，因此，$\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x+\hat{\beta}_2 x^2$回归方程（即曲线的函数表达式）是一个二次多项式，它不是线性的，但可以通过变量代换，化成线性形式。

令$x_1=x$，$x_2=x^2$，原来的回归方程化成了下列形式：

$$\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x_1+\hat{\beta}_2 x_2$$

例2:

在经济学中，有一个著名的科布 – 道格拉斯 (Cobb – Douglas) 生产函数，
这个函数指出，生产产出与劳动投入、资本投入之间，近似有下列关系：

$$Y = \alpha L^{\beta_1} K^{\beta_2}$$

其中，α, β_1, β_2 都是常系数。现测得一组劳动投入、资本投入和生产产出的数据 (L_i, K_i, Y_i)，$i = 1, 2, \cdots, n$，要求从这批数据出发，估计常系数 α, β_1, β_2 的值。

这是一个回归分析问题，回归方程为 $\hat{Y} = \hat{\alpha} L^{\hat{\beta}_1} K^{\hat{\beta}_2}$，它不是线性回归方程，
但是，如果我们对方程两边同时取对数，得到

$$\ln \hat{Y} = \ln \hat{\alpha} + \hat{\beta}_1 \ln L + \hat{\beta}_2 \ln K$$

再令 $y^* = \ln Y$，$\beta_0 = \ln \alpha$，$x_1 = \ln L$，$x_2 = \ln K$，它就化成了一个多元线性回归方程

$$\hat{y}^* = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

例3 在混合异辛烯催化反应中，反应速度y与氢的分压x_1，异辛烯的分压x_2，异辛烷的分压x_3之间，近似有下列关系：

$$y=\frac{kx_1x_2}{(1+a\sqrt{x_1}+bx_2+cx_3)^3}$$

其中，k, a, b, c是常系数，现对x_1, x_2, x_3, y，得到观测值 $(x_{i1}, x_{i2}, x_{i3}, y_i)$，$i=1, 2, \cdots, n$，要求常系数$k, a, b, c$的估计值

对非线性回归方程$\hat{y}=\dfrac{\hat{k}\,x_1x_2}{(1+\hat{a}\sqrt{x_1}+\hat{b}x_2+\hat{c}x_3)^3}$两边开3次方，再取倒数，得到

$$\frac{1}{\sqrt[3]{\hat{y}}}=\frac{1}{\sqrt[3]{\hat{k}\,x_1x_2}}+\frac{\hat{a}\sqrt{x_1}}{\sqrt[3]{\hat{k}\,x_1x_2}}+\frac{\hat{b}\,x_2}{\sqrt[3]{\hat{k}\,x_1x_2}}+\frac{\hat{c}\,x_3}{\sqrt[3]{\hat{k}\,x_1x_2}}$$

$$\Rightarrow \quad \hat{y}^*=\hat{\beta}_1z_1+\hat{\beta}_2z_2+\hat{\beta}_3z_3+\hat{\beta}_4z_4$$

广义线性回归一般形式和解法

问题 设自变量$x_1, x_2, \cdots, x_m$与因变量y之间，有下列关系：

$y = f(\beta_0 + \beta_1\phi_1(x_1, \cdots, x_m) + \cdots + \beta_p\phi_p(x_1, \cdots, x_m)) + \varepsilon$其中，

$y = f(y^*)$是已知的一元函数，有唯一的反函数$y^* = f^{-1}(y)$，

$\phi_1(x_1, \cdots, x_m)$，是自变量$x_1, x_2, \cdots, x_m$的不含未知参数的函数，

$\beta_0, \beta_1, \cdots, \beta_m$是常系数，$\varepsilon \sim N(0, \sigma^2)$是表示误差的随机变量，

$\sigma > 0$，对$x_1, x_2, \cdots, x_m$，y进行n次观测，得到观测值：

$$(x_{i1}, x_{i2}, \cdots, x_{im}, y_i), \quad i = 1, 2, \cdots, n$$

求$\beta_0, \beta_1, \cdots, \beta_m$的估计$\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_m$，使得下式达到最小值：

$$Q = \sum_{i=1}^{n} [\, y_i - f(\beta_0 + \beta_1\phi_1(x_{i1}, \cdots, x_{im}) + \cdots + \beta_p\phi_p(x_{i1}, \cdots, x_{im}))\,]^2$$

分析推导：

对回归方程$\hat{y} = f(\hat{\beta}_0 + \hat{\beta}_1\phi_1(x_1,\cdots,x_m) + \cdots + \hat{\beta}_p\phi_p(x_1,\cdots,x_m))$的两边同时取反函数$f^{-1}$,得到

$$f^{-1}(\hat{y}) = \hat{\beta}_0 + \hat{\beta}_1\phi_1(x_1,\cdots,x_m) + \cdots + \hat{\beta}_p\phi_p(x_1,\cdots,x_m)$$

令$y^* = f^{-1}(y)$，$z_1 = \phi_1(x_1,\cdots,x_m)$，. . . . , $z_p = \phi_p(x_1,\cdots,x_m)$上述方程就化成了线性回归方程$\hat{y}^* = \hat{\beta}_0 + \hat{\beta}_1 z_1 + \hat{\beta}_2 z_2 + \cdots + \hat{\beta}_p z_p$

用线性回归的方法可以求出它的解，得到常系数$\beta_0, \beta_1, \cdots, \beta_m$的估计$\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_m$

广义线性回归中的加权处理

有些广义线性回归问题，化为线性时不需要取反函数$y^* = f^{-1}(y)$，有些则要取反函数$y^* = f^{-1}(y)$ 。对于要取反函数的广义线性回归问题取了反函数后，得到的新问题并不完全等价于原问题。

原问题　　设自变量x与因变量y之间，有下列关系：

$$y = f(\beta_0 + \beta_1 x) + \varepsilon$$

求β_0, β_1的估计$\hat{\beta}_0, \hat{\beta}_1$，使得下式达到最小：

$$Q = \sum_{i=1}^{n} [\, y_i - f(\beta_0 + \beta_1 x_i)\,]^2$$

化为线性后的新问题

在自变量x与因变量y之间的关系式两边取反函数$f^{-1}(y)$，得到

$$Q^* = \sum_{i=1}^{n}[\ f^{-1}(y_i) - (\ \beta_0 + \beta_1 x_i)\ ]^2$$

原来各点到曲线的距离并不等于变换后各点到直线的距离，

所以$\hat{\beta}_0^* \neq \hat{\beta}_0$，$\hat{\beta}_1^* \neq \hat{\beta}_1$

为了解决这一问题，有人提出一种″加权处理″方法

我们知道，当$a \approx b$时，有$f'(a) \approx \dfrac{f(a) - f(b)}{a - b}$

即有$f(a) - f(b) \approx f'(a)(a - b)$

现在因为$f^{-1}(y_i) \approx \beta_0 + \beta_1 x_i$，所以

$$\begin{aligned} y_i - f(\beta_0 + \beta_1 x_i) &= f(f^{-1}(y_i)) - f(\beta_0 + \beta_1 x_i) \\ &\approx f'(f^{-1}(y_i))[f^{-1}(y_i) - (\beta_0 + \beta_1 x_i)] \end{aligned}$$

所以

$$Q=\sum_{i=1}^{n}[\,y_i-f(\,\beta_0+\beta_1 x_i\,)\,]^2\approx\sum_{i=1}^{n}\{f'(f^{-1}(y_i))[f^{-1}(y_i)-(\beta_0+\beta_1 x_i)]\}^2$$

$$=\sum_{i=1}^{n}W_i[f^{-1}(y_i)-(\beta_0+\beta_1 x_i)]^2$$

其中$W_i=[f'(f^{-1}(y_i))]^2$称为权（**Weight**）。

因此，原问题可以近似等价于下列加权回归问题：

求β_0,β_1的估计$\hat{\beta}_0^W,\hat{\beta}_1^W$，使得下式达到最小：

$$Q^W=\sum_{i=1}^{n}W_i[f^{-1}(y_i)-(\beta_0+\beta_1 x_i)]^2$$

由于$Q^W\approx Q$，所以求得的加权最小二乘估计$\hat{\beta}_0^W\approx\hat{\beta}_0$，$\hat{\beta}_1^W\approx\hat{\beta}_1$

例 4 在彩色显影中，形成染料的光学密度 y 与析出银的光学密度 x 之间，近似有下列关系：

$$y = \alpha\, e^{\beta/x} ,$$

其中，α, β 是常系数。现测得数据如下：

x_i	0.05	0.06	0.07	0.10	0.14	0.20	0.25	0.31	0.38	0.43	0.47
y_i	0.10	0.14	0.23	0.37	0.59	0.79	1.00	1.12	1.19	1.25	1.29

求 y 关于 x 的回归方程。

解 对回归方程 $\hat{y} = \hat{\alpha}\, e^{\hat{\beta}/x}$ 两边同时取对数，得到

$$\ln \hat{y} = \ln \hat{\alpha} + \hat{\beta}/x \quad 。$$

令 $y^* = \ln y$ ，$\beta_0 = \ln \alpha$，$\beta_1 = \beta$，$z = 1/x$ ，就把它化成了一个一元线性回归方程

$$\hat{y}^* = \hat{\beta}_0 + \hat{\beta}_1 z \quad 。$$

用 $\ln y_i$ 的值，作为因变量 y^* 观测值 y^*_i，用 $1/x_i$ 的值，作为自变量 z 的观测值 z_i，代入一元线性回归的计算公式，可以求得 $\hat{\beta}_0 = 0.54765$，$\hat{\beta}_1 = -0.14593$，再从下式可以得到原方程中常系数的估计 $\hat{\alpha} = \exp(\hat{\beta}_0) = 1.7292$，$\hat{\beta} = \hat{\beta}_1 = -0.14593$，所以，要求的回归方程为

$$\hat{y} = \hat{\alpha}\, e^{\hat{\beta}/x} = 1.7292\, e^{-0.14593/x} 。$$

它的残差平方和为 $SS_e = 0.0075653$。

以上是在未加权的情况下得到的结果。

在加权的情况下可以求得 $\hat{\beta}_0 = 0.58117$，$\hat{\beta}_1 = -0.15240$，再从下式可以得到原方程中常系数的估计 $\hat{\alpha} = \exp(\hat{\beta}_0) = 1.7881$，$\hat{\beta} = \hat{\beta}_1 = -0.15240$，所以，要求的回归方程为

$$\hat{y} = \hat{\alpha}\, e^{\hat{\beta}/x} = 1.7881\, e^{-0.15240/x} 。$$

它的残差平方和为 $SS_e = 0.0050496$。

与不加权时的情况相比，加权时的残差平方和要小很多，说明加权后得到的回归方程，比起不加权得到的回归方程来，对原始数据拟合得更好，能够更精确地反映自变量与因变量之间的关系。

不能化为线性的非线性回归

问题　设自变量$x_1, x_2, \cdots, x_m$与因变量y之间，有下列关系

$$y = F(x_1, x_2, \cdots, x_m; a_1, a_2, \cdots, a_p) + \varepsilon$$

其中，F是函数形式已知的m元函数，$a_1, a_2, \cdots, a_p$是常数，$\varepsilon \sim N(0, \sigma^2)$

对$x_1, x_2, \cdots, x_m, y$进行n次观测，得到一组观测值：

$$(x_{i1}, x_{i2}, \cdots, x_{im}, y_i), i = 1, 2, \cdots, n$$

求$a_1, a_2, \cdots, a_p$的估计$\hat{a}_1, \hat{a}_2, \cdots, \hat{a}_p$，使得下式达到最小：

$$Q = \sum_{i=1}^{n} [\, y_i - F(\, x_{i1}, x_{i2}, \cdots, x_{im}; a_1, a_2, \cdots, a_p \,)\,]^2$$

分析推导：

$Q = \sum_{i=1}^{n} [\, y_i - F(\, x_{i1}, x_{i2}, \cdots, x_{im}; a_1, a_2, \cdots, a_p \,)\,]^2$是$a_1, a_2, \cdots, a_p$的函数，所以，这是一个多元函数求最小值的问题。

$Q=\sum_{i=1}^{n}[\,y_i-F(\,x_{i\,1},x_{i\,2},\cdots,x_{i\,m}\,;a_1,a_2,\cdots,a_p\,)\,]^2$是$a_1,\,a_2,\cdots,\,a_p$的函数，

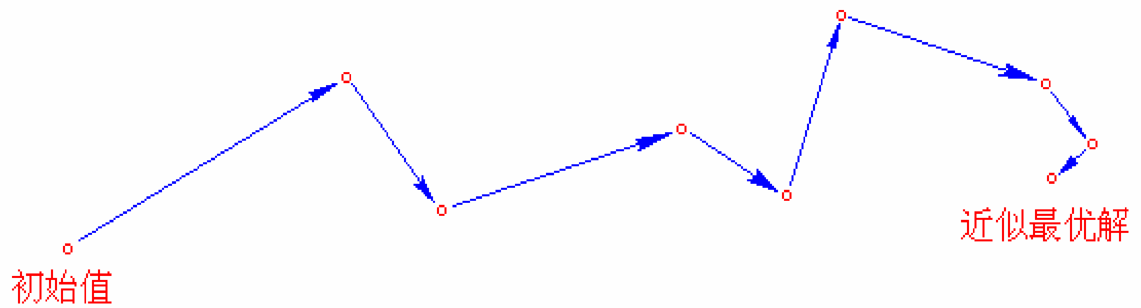
所以，这是一个多元函数求最小值的问题

我们能不能象在线性回归中那样，求的偏导数，令偏导数为 0，
得到一组方程，然后通过解方程组的方法，来确定的最小值点呢？

$$\begin{cases}\dfrac{\partial Q}{\partial a_1}=0\\ \dfrac{\partial Q}{\partial a_2}=0\\ \quad\cdots\\ \dfrac{\partial Q}{\partial a_p}=0\end{cases}$$

从理论上说来，这样做是没有问题的

但是，在许多实际情况下，这样做却是行不通的



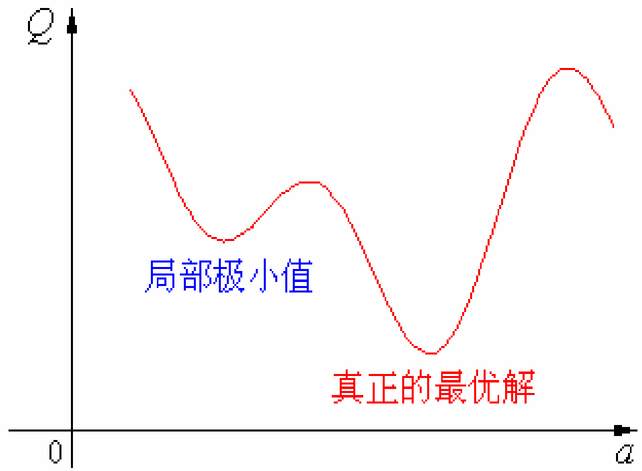
其实，对于这样一个多元函数求最小值的问题，人们已经提出了许多求近似解的数值方法。这些方法的基本思想是：从一个初始值出发，逐步搜索最优解，搜索的步长逐渐缩小，当搜索的步长或最优解的变化小于事先给定的一个误差水平界限时，搜索结束，给出问题的近似解。

在借助计算机软件进行求解的过程中，可能会出现许多问题，例如：

（1）逐步搜索超出函数的定义域，发生计算溢出；

（2）发生“死循环”，长期得不到解；

（3）得到的解不是真正的最优解，等等。



要避免发生这样的问题，选择适当的初始值非常重要。最好能够根据问题的实际意义，参考文献资料和类似的先例，确定比较合理的初始值。还有一种方法，是将问题先作一些简化或转换，变成能用其他回归方法求解的问题，用其他方法求出初步的近似解，再用这个解作为非线性回归的初始值。

例5 对肉鸡的饲养天数 x （单位：日）和肉鸡的重量 y （单位：kg）进行观测，得到一组数据如下：

饲养天数x_i	4	8	12	16	20	24	28	32	36
重量 y_i	0.070	0.119	0.198	0.297	0.434	0.606	0.803	1.027	1.245

饲养天数x_i	40	44	48	52	56	60	64	68
重量 y_i	1.488	1.736	1.980	2.170	2.450	2.687	2.915	3.095

肉鸡重量 y 与饲养天数 x 之间的关系，满足下列微分方程及初始条件：

$$\begin{cases} \dfrac{dy}{dx} = k\,y\,(1-\dfrac{y}{w}) \\ y\big|_{x=0} = y_0 \end{cases}$$

其中，k, w, y_0 是未知常数。求肉鸡重量 y 与饲养天数 x 之间的函数关系。

解 在初始条件下解微分方程，可求得

$$y = \frac{w}{1+(\dfrac{w}{y_0}-1)\mathrm{e}^{-k\,x}}$$ 。

在上式中，常数 w, y_0, k 未知，所以，要得到肉鸡重量 y 与饲养天数 x 之间的关系式，必须从 x 与 y 的观测数据出发，求出 w, y_0, k 的估计值。

这是一个非线性回归问题，回归方程为

$$\hat{y}=\frac{\hat{w}}{1+(\frac{\hat{w}}{\hat{y}_0}-1)\mathrm{e}^{-\hat{k}x}}\quad。$$

利用可以作非线性回归的计算机软件，将数据代入，计算求得 w, y_0, k 的估计值

$$\hat{w}=3.53348\ ,\quad \hat{y}_0=0.106883\ ,\quad \hat{k}=0.0776665\quad。$$

所以，我们得到肉鸡重量 y 与饲养天数 x 之间的函数关系为

$$\hat{y}=\frac{3.53348}{1+(\frac{3.53348}{0.106883}-1)\mathrm{e}^{-0.0776665x}}=\frac{3.53348}{1+32.0593\,\mathrm{e}^{-0.0776665x}}\quad。$$

它的残差平方和为 $SS_e=0.0424689$。这个值相当小，说明回归效果还是很好的。