第四章

经典单方程计量经济学模型

放宽基本假定的模型

基本假设：

1.模型是正确设定的。

4.1 多重共线性

2.X是相互独立的，即$X_j(j=1,2,\ldots,k)$之间互不相关。

3.随机误差项具有给定X条件下的零均值$E(\mu_i|X)=0$，推出$Cov(X_i,\mu_i)=0$。

4.随机误差项具有同方差$Var(\mu_i|X)=\sigma^2$、不序列相关$Cov(\mu_i,\mu_j|X)=0$。

4.2 异方差性

5.随机误差项服从正态分布。

第四章 放宽基本假定的模型

4.1 多重共线性

4.2 异方差性

4.3 内生解释变量问题

4.4 模型设定偏误问题

4.5 序列相关性

4.2 异方差性

一、异方差性的概念和类型

对于模型

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \mu_i$$

同方差假设为

$$Var(\mu_i | X_{1i}, X_{2i}, \ldots, X_{ki}) = \sigma^2$$

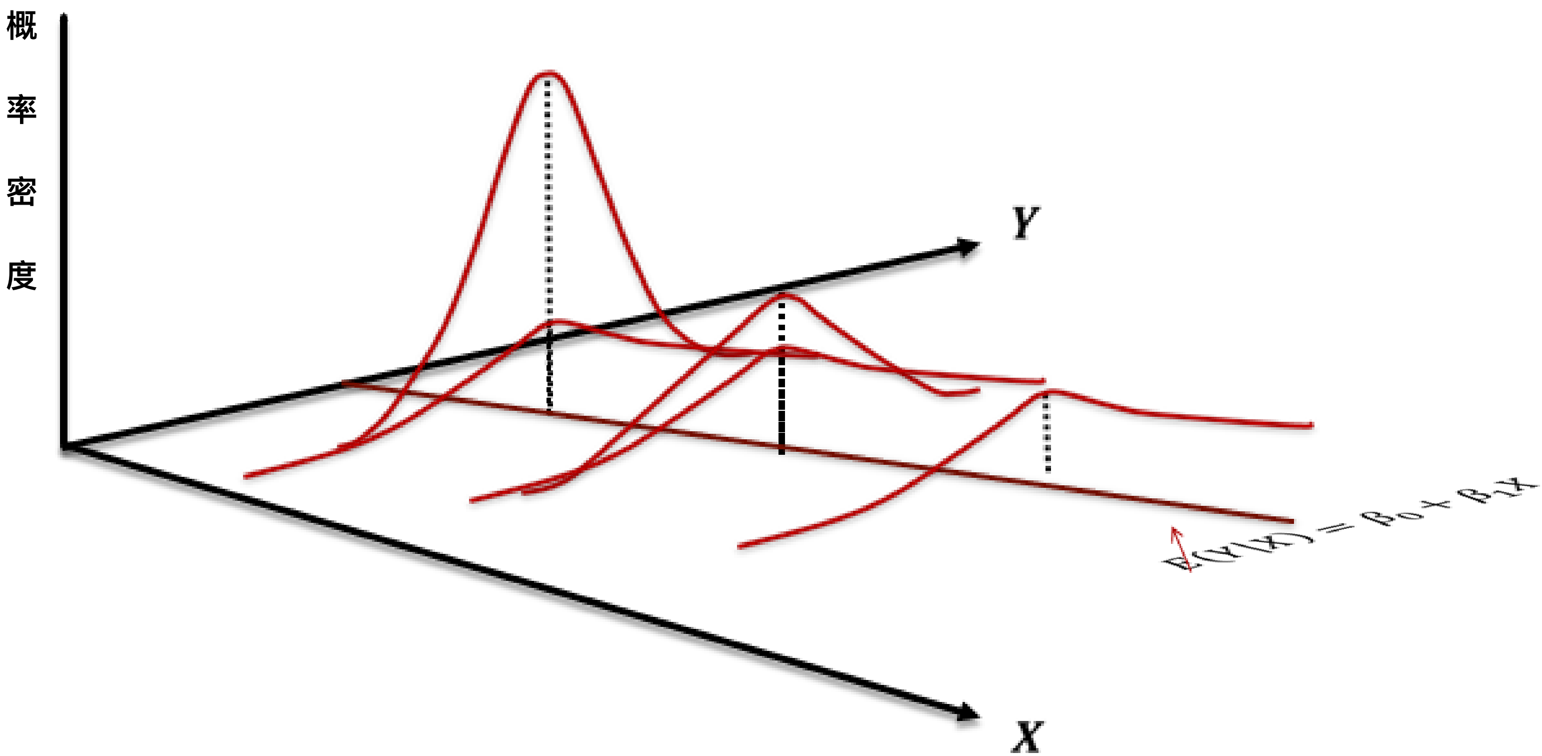
如果出现

$$Var(\mu_i | X_{1i}, X_{2i}, \ldots, X_{ki}) = \sigma_i^2$$

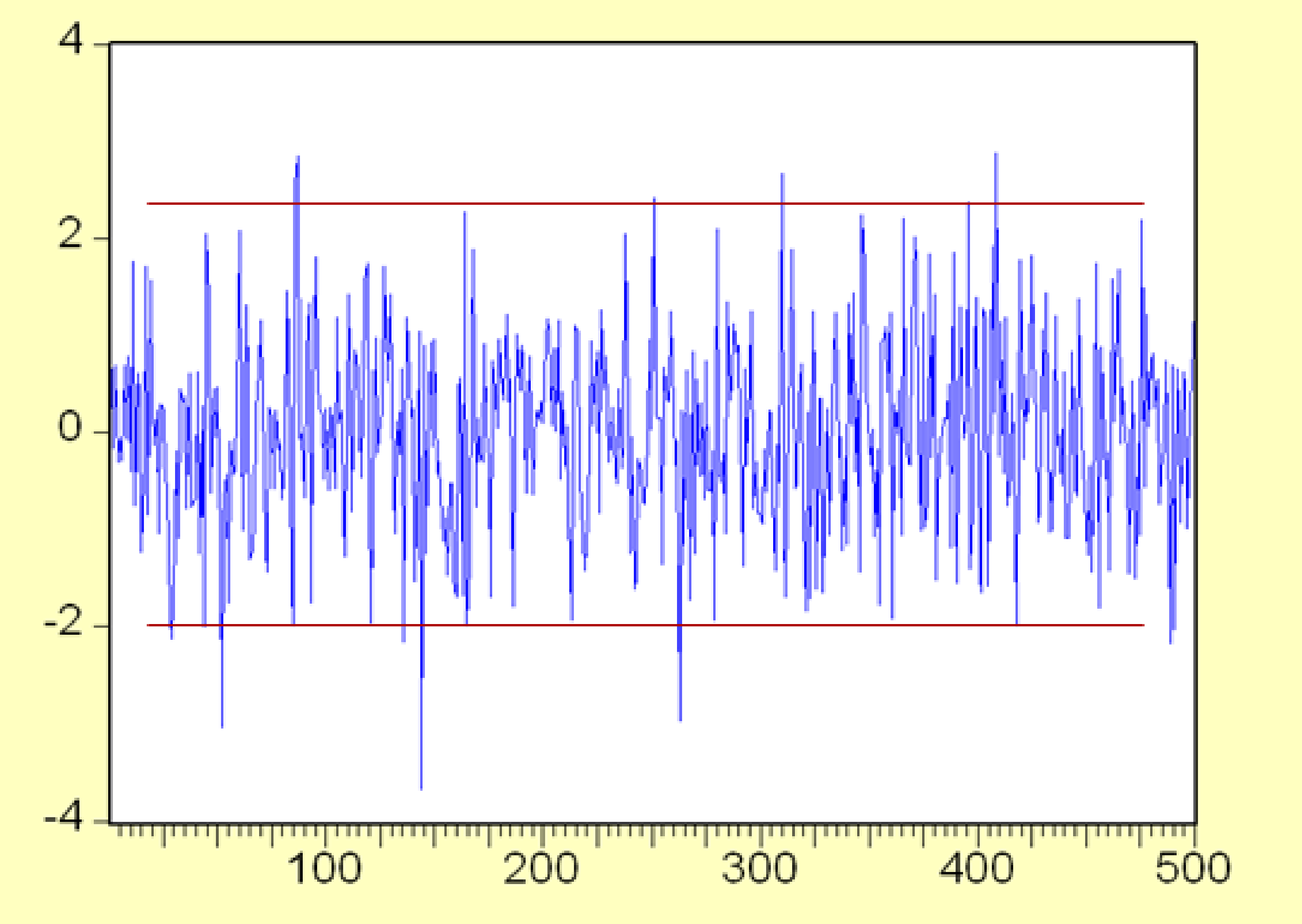
即**对于不同样本点，随机干扰项的方差不再是常数**，而是互不相同，则认为出现了异方差性（heteroscedasticity）。

方差度量的是被解释变量观测值围绕回归线的分散程度。

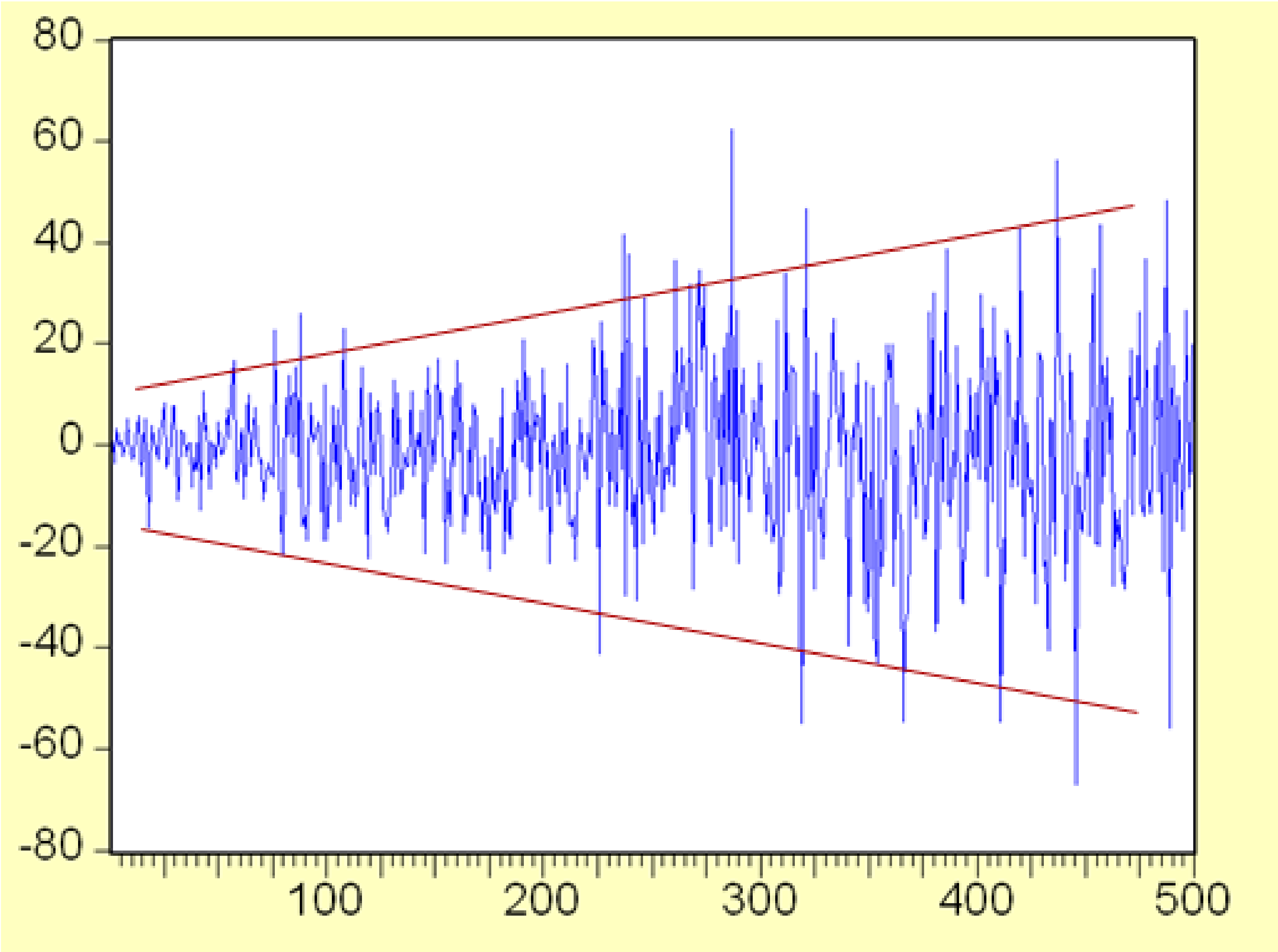
异方差性就是指被解释变量观测值的分散程度随解释变量的变化而变化。



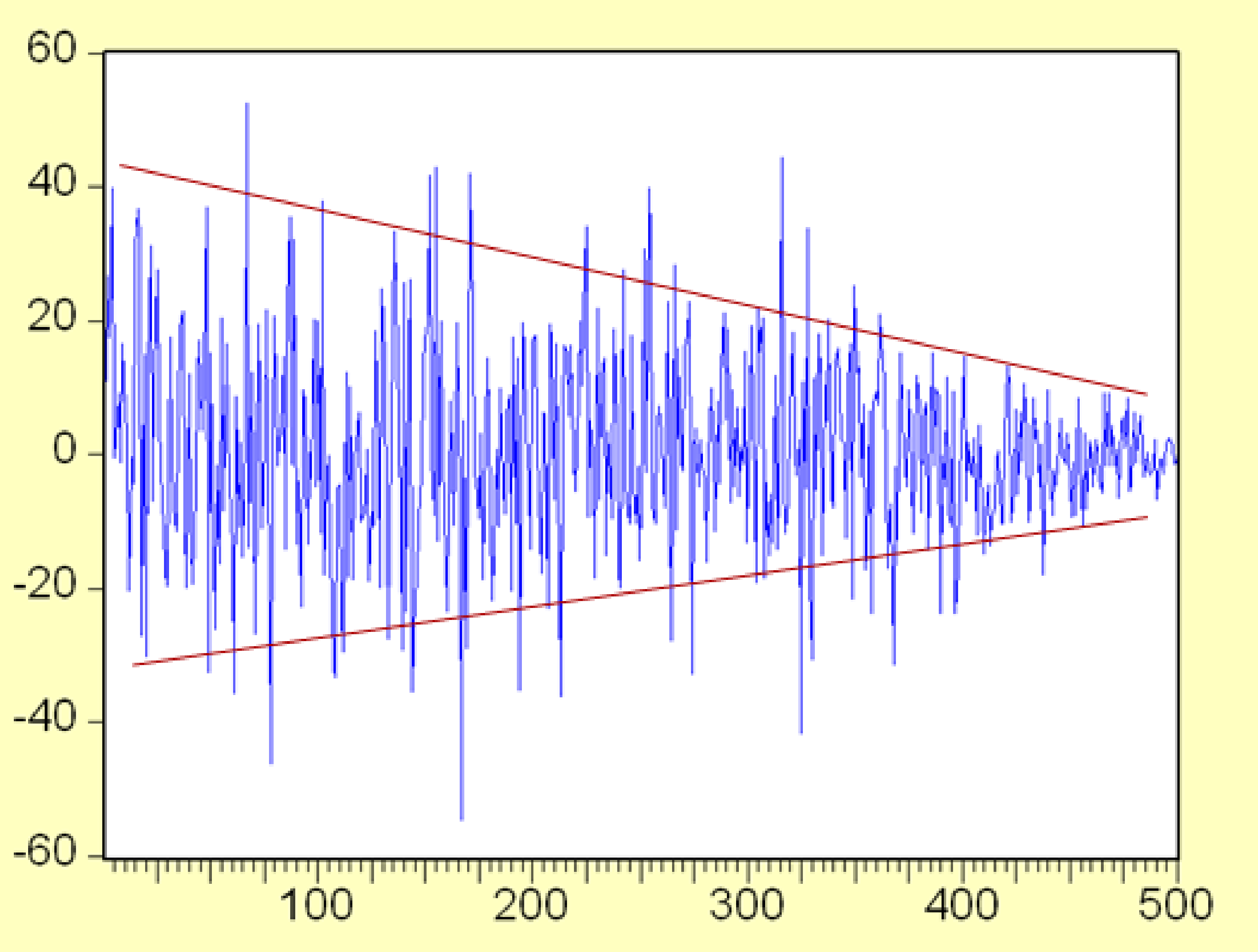
同方差序列（线性图）



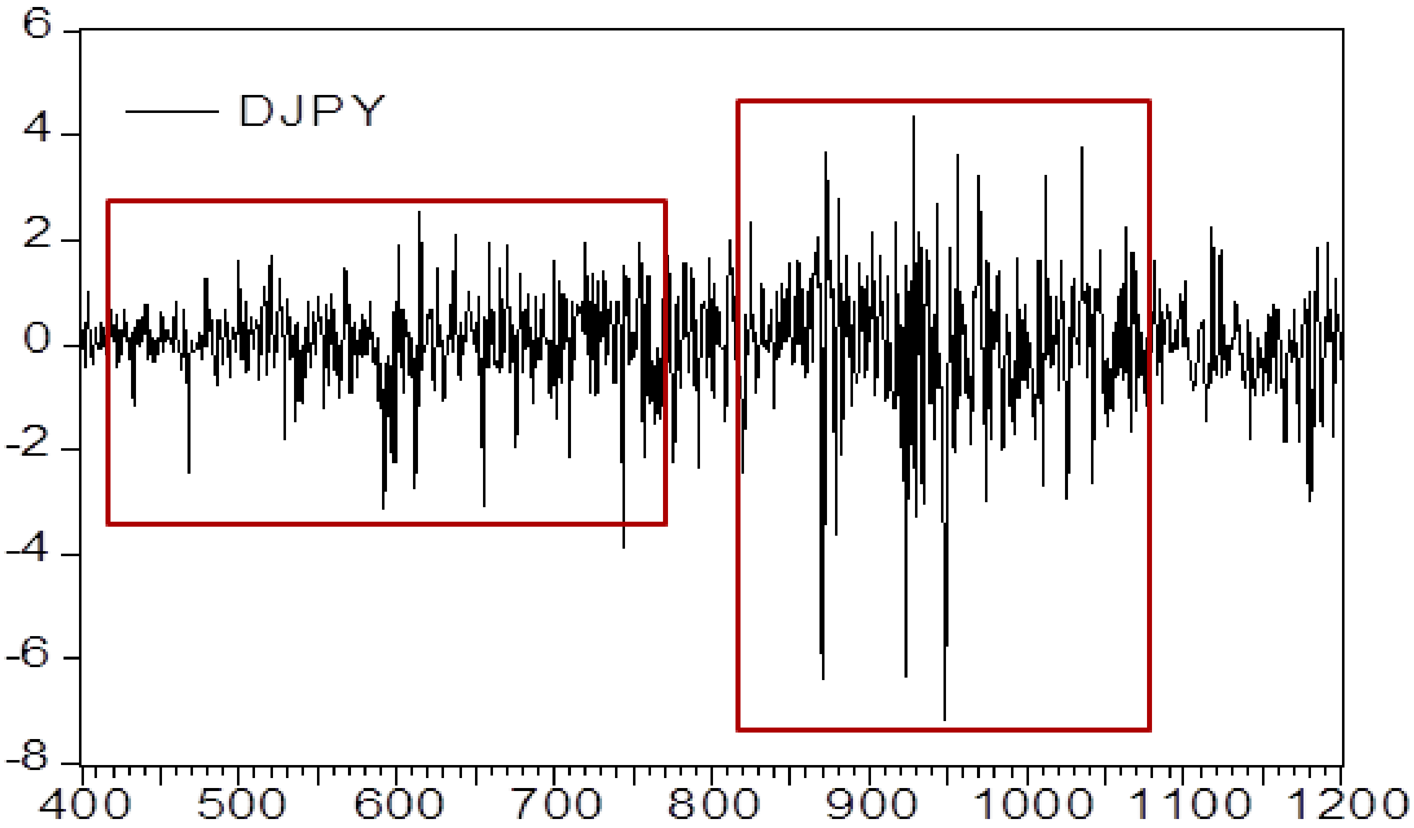
方差逐渐增大的序列（线性图）



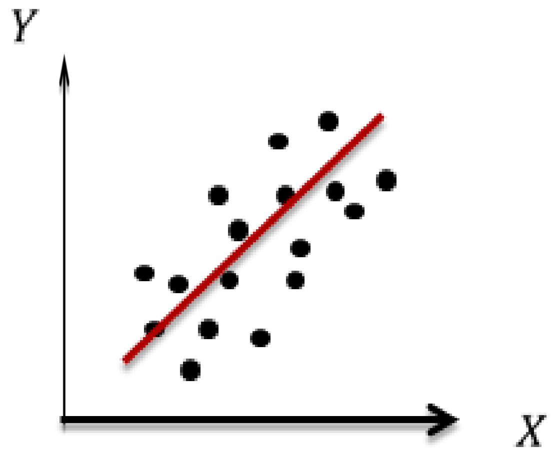
方差逐渐减小的序列（线性图）



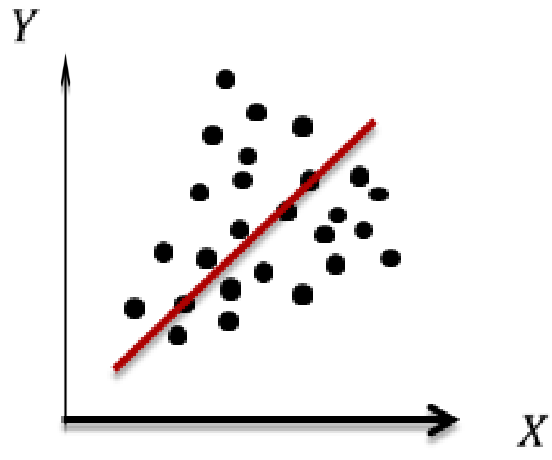
复杂型异方差序列（线性图）



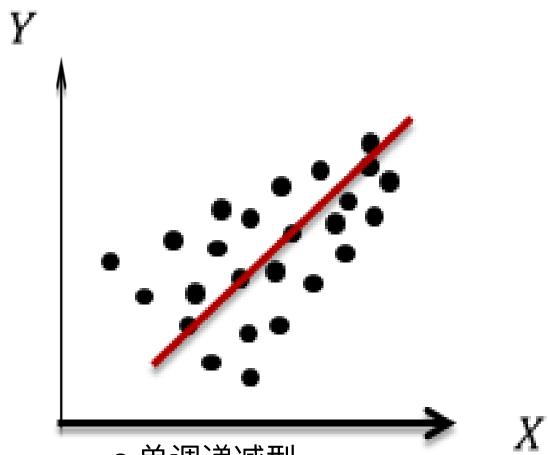
异方差的类型图示（散点图）



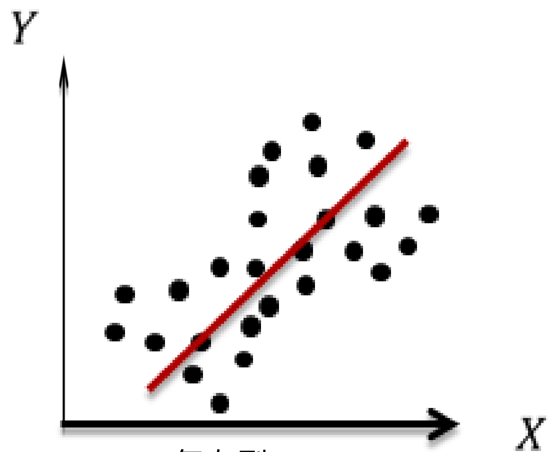
a.同方差



b.单调递增型



c.单调递减型



d.复杂型

二、产生异方差问题的原因

（一）截面数据中总体各单位的差异

【例】研究居民的储蓄行为（横截面数据）

$$S_i = \beta_0 + \beta_1 Y_i + \mu_i$$

S_i —— 第i个家庭的储蓄额

Y_i —— 第i个家庭的可支配收入

- 对低收入家庭来说，储蓄有规律性（如为某一特定目的而储蓄），差异较小；
- 对高收入家庭来说，储蓄的差异较大。

因此，μ_i的方差随着X_i的增大而增加，呈现单调递增型变化。

【例】建立企业生产函数模型（以某一行业内的各企业为样本）

$$Y_i = \beta_0 A_i^{\beta_1} K_i^{\beta_2} L_i^{\beta_3} e^{\mu_i}$$

$$\begin{cases} \text{被解释变量——产出量}Y_i \\ \text{解释变量——资本}K\text{、劳动}L\text{、技术}A \end{cases}$$

- 每个企业所处的外部环境对产出量的影响被包含在随机误差项中。
- 每个企业所处的外部环境对产出量的影响程度不同，造就了随机干扰项的异方差性。
- 随机误差项的方差并不随某一个解释变量观测值的变化呈规律性变化，而呈现复杂型变化。

（二）模型设定误差

选择错误的变量：如模型正确形式为$Y_i=\beta_0+\beta_1X_{i1}+\beta_2X_{i2}+\beta_3X_{i3}+\mu_i$，若忽略了$X_3$，会导致$X_3$对$Y$的影响反应在$\mu$中，而这些影响具有差异性。所以在用剔除法去除多重共线性时，又有可能引起异方差性，应该注意。

选择错误的函数形式：如把非线性关系设定为线性，可导致异方差产生。

（三）数据的观测误差

【例】建立居民消费函数（横截面数据）

$$C_i = \beta_0 + \beta_1 Y_i + \mu_i$$

将居民按收入分成n组，取组平均数为样本观测值。

一般情况下，居民收入服从正态分布，中等收入组人数多，平均数的误差小；两端收入组人数少，平均数误差大。

由于样本观测值的观测误差随着解释变量观测值的不同而不同，因而导致异方差性。

三、异方差性的后果

计量经济学模型一旦出现异方差性，如果仍采用OLS估计模型参数，会产生下列不良后果：

1、参数估计量无偏但非有效；

2、变量的显著性检验失去意义；

3、模型的预测失效。

（一）参数估计量无偏但非有效

● 估计量无偏性的证明

在解释变量样本值给定的条件下，参数估计量$\hat{\beta}$具有无偏性。

$$
\begin{aligned}
E(\hat{\beta}|X) &= E\left(\left(X'X\right)^{-1}X'Y|X\right) \\
&= E\left(\left(X'X\right)^{-1}X'(X\beta+\mu)|X\right) \\
&= \beta+(X'X)^{-1}X'E(\mu|X) \\
&= \beta
\end{aligned}
$$

这里利用了随机干扰项条件零均值假设$E(\mu|X)=0$。

● 估计量有效性的证明

参数β的方差-协方差矩阵可进行如下转变

$$
\begin{aligned}
Var-Cov(\hat{\beta}) &= E[\left(\hat{\beta}-E(\hat{\beta})\right)\left(\hat{\beta}-E(\hat{\beta})\right)'|X] \\
&= E[(\hat{\beta}-\beta)(\hat{\beta}-\beta)'|X] \\
&= E[(X'X)^{-1}X'\mu\mu'X(X'X)^{-1}|X] \\
&= (X'X)^{-1}X'E(\mu\mu'|X)X(X'X)^{-1} \\
&= (X'X)^{-1}X'\sigma^2 I_n X(X'X)^{-1} \\
&= \sigma^2(X'X)^{-1}
\end{aligned}
$$

$$
\boxed{
\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'Y \\
&= (X'X)^{-1}X'(X\beta+\mu) \\
&= \beta+(X'X)^{-1}X'\mu
\end{aligned}
}
$$

因为有效性证明中利用了$E(\mu\mu'|X)=\sigma^2 I_n$，而存在异方差时$\sigma^2$发生了改变，因而估计量不具有效性。

（二）变量的显著性检验失去意义

t统计量是建立在随机干扰项共同的方差σ^2不变，而正确估计了参数方差$Se(\hat{\beta}_j)$的基础之上的

$$t=\frac{\hat{\beta}_j-\beta_j}{Se(\hat{\beta}_j)}=\frac{\hat{\beta}_j-\beta_j}{\sqrt{\sigma^2(X'X)^{-1}}}$$

当异方差性出现时，σ^2发生改变，$Se(\hat{\beta}_j)$出现偏误（偏大或偏小）， t检验因此失去意义。

（三）模型预测失效

- 一方面，由于上述后果，使得模型不具有良好的统计性质。
- 另一方面，当模型出现异方差性时，参数OLS估计值的变异程度增大，从而造成对Y的预测误差变大，降低预测精度，预测功能失效。

视频片段

四、异方差性的检验

检验思路：异方差是指对于不同的[redacted]

随机扰动项具有不同的[redacted]

检验就是检验随机扰动[redacted]

相关性及相关形式。

那么应该如何正确表示$Var(\mu)$呢？

一般是采用OLS方法估计模型，求得μ的估计量，称为“近似估计量”，记做e_i。

$$e_i = Y_i - (\hat{Y}_i)_{OLS}$$

异方差性的检验方法：

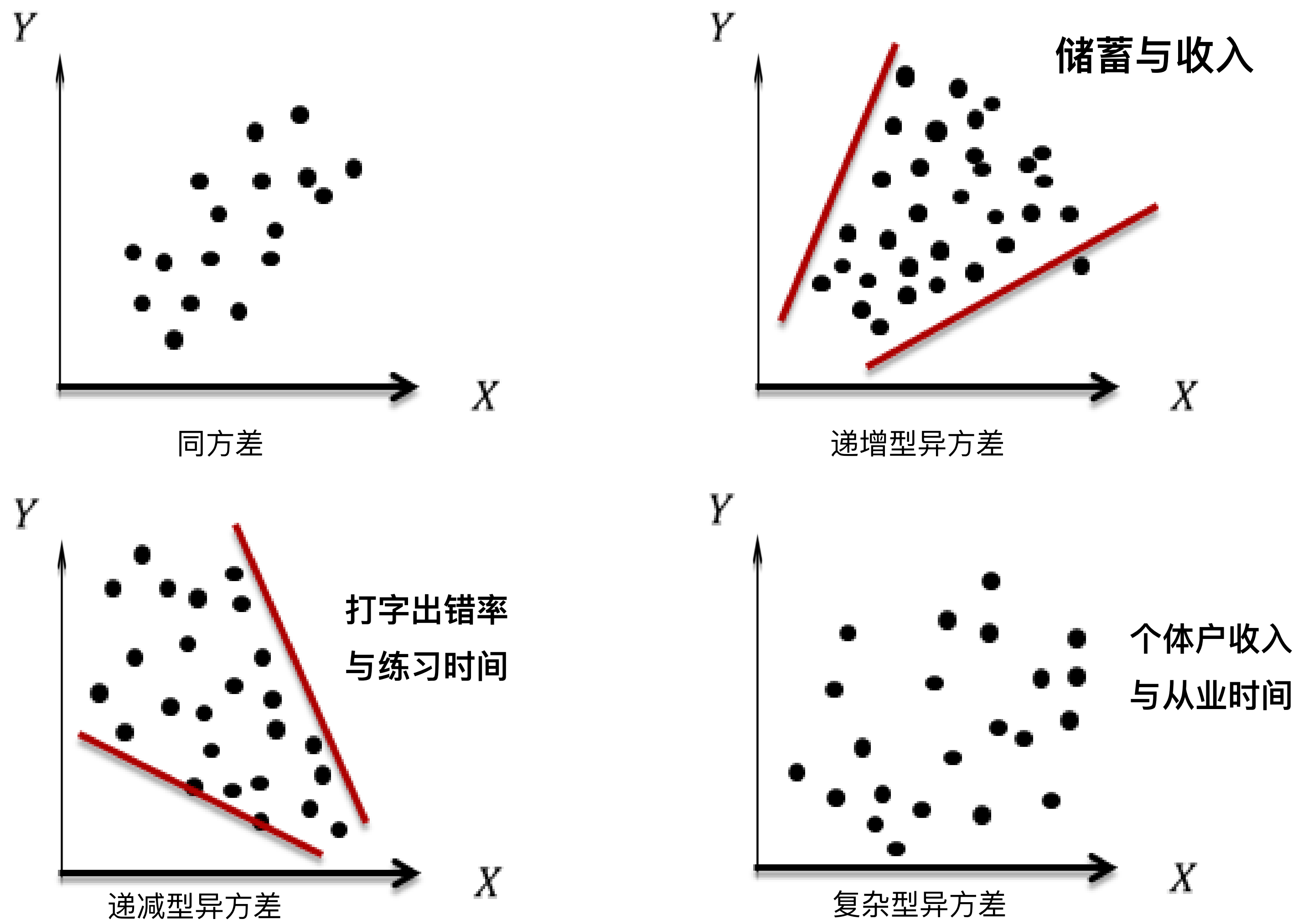
（一）图示检验法

① 用$Y-X$散点图判断

② 用e_i^2-X散点图判断

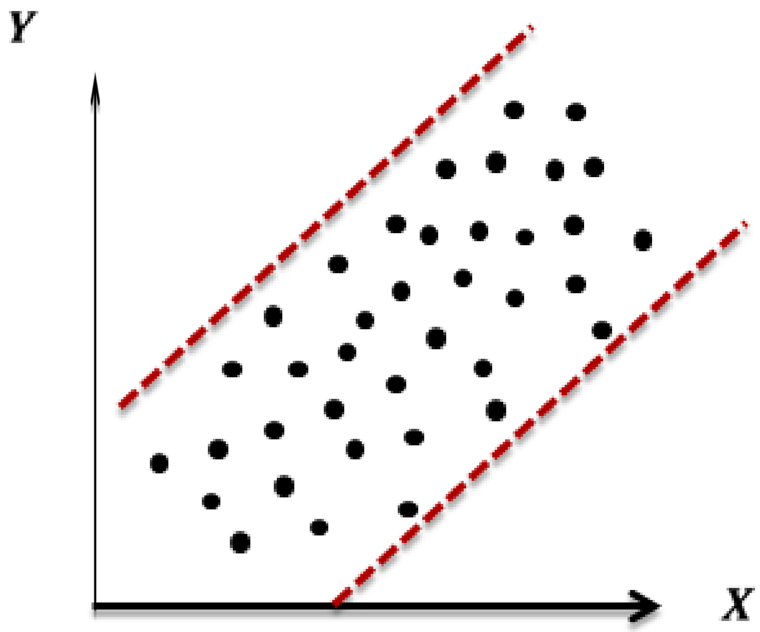
（二）怀特检验

（一）图示检验法—— $Y-X$散点图

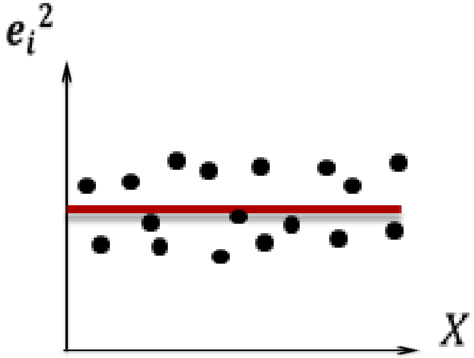


$Y-X$散点图判断异方差性的原则：

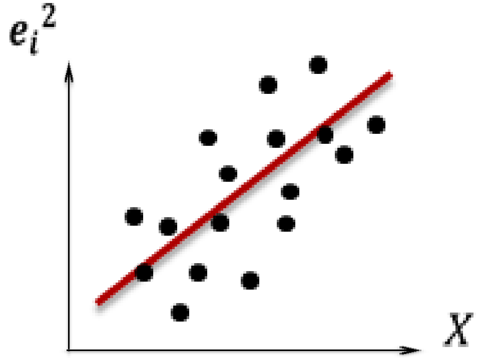
利用$Y-X$进行异方差性判断，要观察是否存在明显的散点扩大、缩小或复杂性趋势，即**散点是否在固定带域中。**



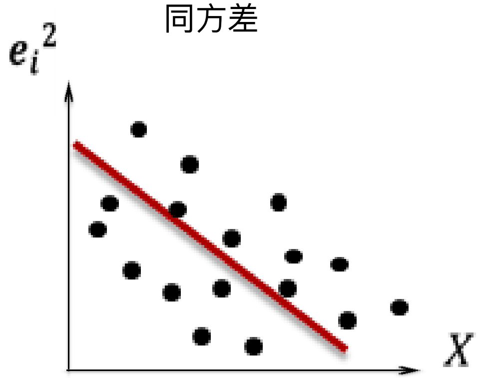
（一）图示检验法—— e_i^2-X散点图



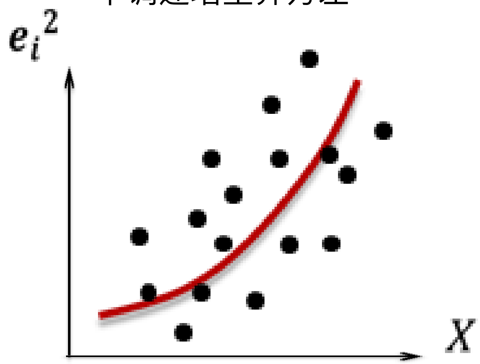
同方差



单调递增型异方差



单调递减型异方差



复杂型异方差

e_i^2-X散点图判断异方差性的原则：是否形成一条斜率为0的直线。

（二）怀特（White）检验

怀特检验的优点：适用于任何形式的异方差检验。

检验思路：异方差性意味着μ^2是部分或者全部解释变量的某种函数，因此这种函数可以是非线性的，既可以包含解释变量的平方项以及不同解释变量之间的交叉项。

【注意】

同方差假设意味着

$$Var(\mu_i|X_{1i}, X_{2i}, \dots, X_{ki}) = \sigma^2$$

随机干扰项服从条件零均值时

$$Var(\mu_i|X_i) = E(\mu_i^2|X_i) - (E(\mu_i|X_i))^2 = E(\mu_i^2) \approx {e_i}^2$$

● 怀特检验基本思想（以二元为例）：

对模型

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \mu_i$$

进行OLS回归，得到μ_i^2的近似估计量${e_i}^2$

做辅助回归

$${e_i}^2 = \alpha_0 + \alpha_1 X_{i1} + \alpha_2 X_{i2} + \alpha_3 {X_{i1}}^2 + \alpha_4 {X_{i2}}^2 + \alpha_5 X_{i1} X_{i2} + \varepsilon_i$$

要检验的同方差性假设为

$$H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$$

在同方差假设下，可以通过两个统计量考察是否存在异方差

拉格朗日乘数(LM)统计量：$LM = nR^2_{e^2} \dot{\sim} \chi^2(k)$

F统计量为 ：$F = \dfrac{R^2_{e^2}/k}{\left(1-R^2_{e^2}\right)/(n-k-1)} \sim F\ (\ k, n-k-1\ )$

其中：n—样本容量，$R^2_{e^2}$—辅助回归的可决系数，k—辅助回归解释变量个数。

判定准则：如果计算出的F值或LM值大于给定显著性水平下的临界值，则拒绝H_0，表明存在异方差性。

【注意】辅助回归可选择解释变量的更高次方。也可以去掉交叉项或平方项。

● 怀特检验具体步骤（利用LM 统计量）

<1> 确定原假设和备择假设

H_0：μ_i为同方差　　　H_1：μ_i为异方差

<2> 做出辅助回归，计算$nR_{e^2}^2$

<3> 给定α，查表得到$\chi_\alpha{}^2(k)$

<4> 比较，判断

若$nR_{e^2}^2 \leq \chi^2(k)$，接受$H_0$，$\mu_i$为同方差。

若$nR_{e^2}^2 > \chi^2(k)$，接受H_1，μ_i为异方差。

【注意】辅助回归判定系数越大，解释变量受e_i影响越大。

五、异方差性的克服——广义（加权）最小二乘法

（一）广义最小二乘法的原理：

广义（加权）最小二乘法是对原模型进行加权，使之成为一个不存在异方差性的新模型，然后利用OLS对模型进行估计。

（二）广义最小二乘法基本思想：

在OLS下，对较小的e_i赋予较大的权数，对较大的e_i赋予较小的权数，即对残差提供的信息的重要程度做出校正，提高估计精度。

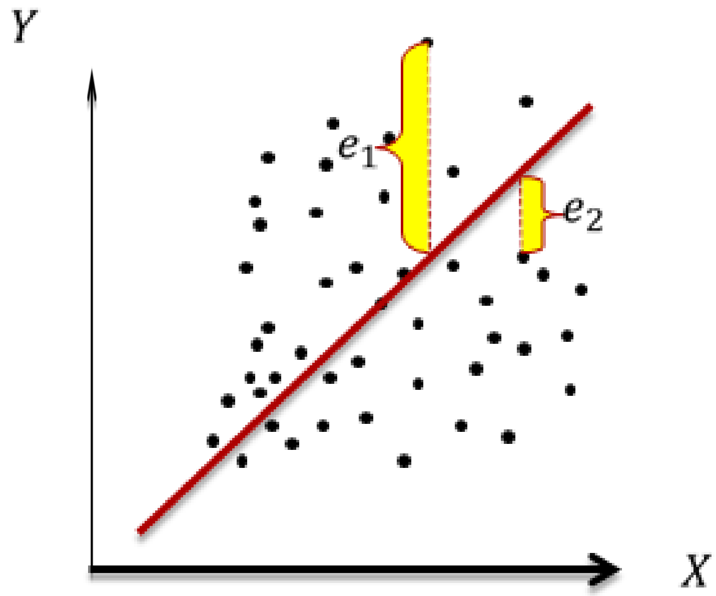
加权最小二乘法即对加了权重的残差平方和实施OLS

$$\sum w_i {e_i}^2 = \sum w_i \left[Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_k X_k)\right]^2$$

其中，w_i为权数。

给紧密围绕（总体）均值的观测值较大权数，给远离均值的观测值较小的权数

——古扎拉蒂



（三）加权最小二乘法中权数的选择

选择权重，应该根据是否已经知道μ的方差与解释变量的具体相关形式，可以将权数的选择以及估计方法分为两种：

① μ的方差与各X间关系为已知的函数形式——**加权最小二乘估计法**

② μ的方差与各X间关系为未知的函数形式——**可行的加权最小二乘法**

1. 加权最小二乘估计法

对于多元回归模型

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \mu_i$$

假定

$$\mathrm{Var}(\mu_i) = \mathrm{E}(\mu_i^2) = \sigma_i^2 = f(X_{ij})\sigma^2$$

其中$f(X_{ij}) > 0$是某些导致异方差性产生的解释变量观测值的函数，该式表明μ_i的方差与某些解释变量之间存在相关性

在$f(X_{ij})$已知的情况下，用$\sqrt{f(X_{ij})}$去除原模型，得到新模型

$$\frac{1}{\sqrt{f(X_{ij})}}Y_i = \beta_0 \frac{1}{\sqrt{f(X_{ij})}} + \beta_1 \frac{1}{\sqrt{f(X_{ij})}} X_{i1} + \cdots + \beta_k \frac{1}{\sqrt{f(X_{ij})}} X_{ik} + \frac{1}{\sqrt{f(X_{ij})}} \mu_i$$

新模型中随机扰动项的方差为

$$Var\left(\frac{1}{\sqrt{f(X_{ij})}}\mu_i\right) = \left(\frac{1}{\sqrt{f(X_{ij})}}\right)^2 Var(\mu_i)$$

$$= \frac{1}{f(X_{ij})} f(X_{ij}) \sigma^2$$

$$= \sigma^2$$

满足同方差性，可以利用OLS估计。

其中，权数为$w_i = \dfrac{1}{\sqrt{f(X_{ij})}}$。

如果发现

$$Var(\mu_i|X_{i1}, X_{i2}, \dots, X_{ik}) = \sigma^2 f(X_{i1}, X_{i2}, \dots, X_{ik})$$

则加权最小二乘法中的权即为

$$w_i = 1/\sqrt{f(X_{i1}, X_{i2}, \dots, X_{ik})}$$

2. 可行的加权最小二乘估计法

定义： 当$f(X_{ij})$未知时，把估计的$f(X_{ij})$记为$\hat{f}(X_{ij})$，用$\hat{f}(X_{ij})$作为权重函数得到估计量，称为可行的广义最小二乘估计。

估计步骤：

假定μ的方差具有如下的指数函数形式

$$Var(\mu_i) = E(\mu_i{}^2) = \sigma^2 e^{\alpha_0+\alpha_1 X_{i1}+\alpha_2 X_{i2}+\cdots+\alpha_k X_{ik}}$$

其等价于

$$\mu_i{}^2 = \sigma^2 e^{(\alpha_0+\alpha_1 X_{i1}+\alpha_2 X_{i2}+\cdots+\alpha_k X_{ik})} \varepsilon_i$$

其中，ε_i为条件均值为1的随机项。

取对数得

$$ln(\mu_i{}^2) = \delta_0 + \alpha_1 X_{i1} + \alpha_2 X_{i2} + \cdots + \alpha_k X_{ik} + v_i$$

其中： v_i为独立于各X，且条件均值为0的随机项。

用$e_i{}^2$代替$\mu_i{}^2$，得到

$$ln(e_i{}^2) = \delta_0 + \alpha_1 X_{i1} + \alpha_2 X_{i2} + \cdots + \alpha_k X_{ik} + v_i$$

满足基本假设，可以估计各参数$\hat{\alpha}_j (j = 1, 2, \ldots, k)$

从而得到μ的方差估计

$$\hat{\sigma}_i{}^2 = \hat{\mu}_i{}^2 = \hat{f}(X_{ij}) = e^{(\hat{\delta}_0 + \hat{\alpha}_1 X_{i1} + \hat{\alpha}_2 X_{i2} + \cdots + \hat{\alpha}_k X_{ik})}$$

此时，估计的权数为

$$w_i=\frac{\mathbf{1}}{\hat{\sigma}_i}=\frac{1}{\sqrt{\hat{f}(X_{ij})}}=\frac{1}{\sqrt{e^{(\hat{\delta}_0+\hat{\alpha}_1X_{i1}+\hat{\alpha}_2X_{i2}+\cdots+\hat{\alpha}_kX_{ik})}}}$$

【注意】

在

$$Var(\mu_i)=E(\mu_i{}^2)=\sigma^2e^{\alpha_0+\alpha_1X_{i1}+\alpha_2X_{i2}+\cdots+\alpha_kX_{ik}}$$

中，只列出了各解释变量X的水平项，可根据估计值的显著性，对各X进行取舍。并且还可以根据需要加入适当的X的高次方项。

视频片段