

$$Y_i = \beta_1 + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \mu_i \quad i = 1, 2, \dots, n$$

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \mu_i, \text{其中 } X_{1i} = 1$$

如果某两个或多个解释变量之间出现了相关性，则称为**多重共线性(Multicollinearity)**。

回归模型中解释变量的关系可能为以下三种：

- (1)  $r_{X_i X_j} = 0$ ，解释变量间无线性关系，变量正交，每个参数都可用Y与X<sub>j</sub>的一元线性回归来估计；
- (2)  $r_{X_i X_j} = 1$ ，解释变量完全共线，模型参数无法估计；
- (3)  $0 < r_{X_i X_j} < 1$ ，解释变量之间存在一定的线性关系，实际中常遇到，共线程度越强，参数估计的准确性稳定性受影响越大。

## 1、简单相关系数检验法

- 对两个解释变量的模型，采用**简单相关系数法**。
  - 求出 $X_1$ 与 $X_2$ 的简单相关系数 $r$ ，若 $|r|$ 接近1，则说明两变量存在较强的多重共线性。
- 对多个解释变量的模型，采用**综合统计检验法**。
  - 如果在**OLS法**下， $R^2$ 与F值较大，但t检验值较小，说明各解释变量对Y的联合线性作用显著，但各解释变量间存在共线性而使得它们对Y的独立作用不能分辨，故t检验不显著。

$R^2 \uparrow$   $F \uparrow$   $t \downarrow$

## 2、方差膨胀因子法

- 使模型中每一个解释变量分别以其余解释变量为解释变量进行**辅助回归 (Auxiliary Regression)**，并计算相应的拟合优度。
- 如果某一种回归 $X_{ji} = \alpha_1 X_{1i} + \alpha_2 X_{2i} + \dots + \alpha_L X_{Li}$ 的 $R^2$ 较大，说明 $X_j$ 与其他X间存在**共线性**。
- 也可以直接计算解释变量的方差膨胀因子。

$$VIF = \frac{1}{R^2}$$

越大越强

### 3、直观判断法

- 在模型中排除某一个解释变量 $X_j$ ，估计模型；如果拟合优度与包含 $X_j$ 时十分接近，则说明 $X_j$ 与其它解释变量之间存在共线性。
- 重要的解释变量在t检验中不显著时，可能存在严重的多重共线。
- 解释变量的回归系数正负号与定性分析相违背时。
- 如果在OLS法下， $R^2$ 与F值较大，但t检验值较小，说明各解释变量对Y的联合线性作用显著，但各解释变量间存在共线性而使得它们对Y的独立作用不能分辨，故t检验不显著。

### 4. 逐步回归法 (Stepwise forward Regression)

- 以Y为被解释变量，逐个引入解释变量，构成回归模型，进行模型估计。
- 根据拟合优度的变化决定新引入的变量是否独立。
  - 如果拟合优度变化显著，则说明新引入的变量是一个独立解释变量；
  - 如果拟合优度变化很不显著，则说明新引入的变量与其它变量之间存在共线性关系。

## 一、修正多重共线性的经验方法

### 1.剔除变量法

找出引起多重共线性的解释变量，将它排除。

- 以逐步回归法得到最广泛的应用。
- 注意：剩余解释变量参数的经济含义和数值都发生了变化。

### 2.增大样本容量

可使参数估计量的方差减小

### 3.变换模型形式

如差分法

### 4.利用非样本先验信息

通过经济理论分析能够得到某些参数之间的线性关系，可将这种线性关系作为约束条件，与样本信息结合起来做。

### 5.横截面数据与时序数据并用

## 6. 变量变换

- (1) 计算相对指标
- (2) 名义数据转换为实际数据
- (3) 小类指标合成大类指标
- (4) 对总量指标进行对数变换

变量变换只是一种降低多重共线的方法，不一定能得到好效果。

