放宽条件的模型汇总

概念	背景	后果	检验	修正
多重共线性：X间高度相关 =完全~、近似~	模型设定失误 经济变量存在共同趋势 样本资料限制 （当前条件下无法找到满足不相关的样本数据）	破坏经典假设2 ⇒破坏最小方差性 ⇒t检验失效、预测失效 参数估计的经济意义不合理 完全共线性下参数估计不存在	存在 简单相关系数法 单元，计算两个变量的相关系数r 综合统计检验法 多元，如果R^2和F比较大，t比较小 **范围** 判定系数检验法 对每一个解释变量，用其他解释变量进行回归，F检验 逐步回归法 检验&解决，逐个引入变量，如果R^2显著改善，引入，否则不 方差膨胀因子 $VIF = 1/(1-R^2)$	逐步回归法
异方差性：μ的方差不同		破坏经典假设4 ⇒破坏最小方差性 ⇒t检验失效、预测失效	图示法（注意区分纵坐标） $X-Y$散点图 $X-\tilde{e}_i^2$散点图 B-P检验 [1] 残差可以由解释变量线性表示 用X对e_i^2回归 LM统计量 White检验 [2] 第一步：OLS得到e_i 第二步：用二次项对e_i^2进行回归 LM统计量	WLS加权最小二乘 [3] 异方差稳健标准误 [4]
内生解释变量：X和μ相关 同期相关、异期相关	联立因果、互为因果 遗漏相关解释变量 解释变量存在测量误差	破坏无偏性 [5] （只有这个不是破坏最小方差性！） ⇒OLS估计量	Hausman检验 [6] 第一步：对内生解释变量进行回归，取残差 第二步：将残差加入原模型	工具变量法 [8] 相关、外生、与其他解释变量不高度相关

		失效	过度识别的约束检验[7] 第一步：2SLS，记录残差项 第二步：用所有工具变量及外生变量进行回归 LM统计量	2SLS[9]：过度识别 第一步：估计内生解释变量 第二步：用$\hat{X}$替换X进行回归
模型设定偏误 遗漏解释变量（如果遗漏了相关的X，就是内生性） 误选无关变量（无偏，但影响最小方差性） 错误的函数形式			残差图示法 RESET检验 先回归，然后将$\hat{Y}$的若干次幂引入模型重新估计	
序列相关性：$\mathrm{Cov}(\mu_i, \mu_j) \neq 0$ 一阶序列相关：自协方差系数ρ	经济变量固有惯性 模型设定偏误 数据编造	破坏最小方差性 $\Rightarrow t$检验失效、预测失效	图示法 注意横坐标是t还是$\tilde{e}_{t-1}$ 回归检验法[10] 直接检验 杜宾DW检验[11] 检验一阶自相关 拉格朗日乘数检验[12] 检验高阶序列相关	GLS[13] 广义差分法[14] 消随机误差项 科克伦-奥科特迭代法 随机误差项相关系数的估计 先原模型算μ，给出ρ 再广义差分模型算β

1. $e_i^2 = \delta_0 + \delta_1 X_{i1} + \cdots + \delta_k X_{ik} + \varepsilon_i$，$H_0: \delta_1 = \cdots = \delta_k = 0$，$LM = n \cdot R_{e^2}^2 \dot{\sim} \chi^2(k)$ ↩
2. $e_i^2 = \alpha_0 + \alpha_1 X_{i1} + \alpha_2 X_{i2} + \alpha_3 X_{i1}^2 + \alpha_4 X_{i2}^2 + \alpha_5 X_{i1} X_{i2} + \varepsilon_i$，$H_0: \alpha_1 = \cdots = \alpha_5 = 0$，$LM = nR^2 \dot{\sim} \chi^2(5)$ ↩
3. 如果发现$\mathrm{Var}(\mu_i) = \sigma_i^2 = f(X_{ij})\sigma^2$，就在模型两遍同时除以$f$，假设方差存在指数函数形式$\mu_i^2 = \sigma^2 \exp(\alpha_0 + \alpha_1 X_{i1} + \cdots + \alpha_k X_{ik})$，对模型$\ln(\tilde{e}_i^2) = \delta_0 + \alpha_1 X_{i1} + \cdots + \alpha_k X_{ik} + \nu_i$进行估计 ↩
4. 为异方差满足无偏性和一致性（参数估计没问题），只需要修正方差。因此，用e_i^2代替σ_i^2（替换后满足大样本下一致性） ↩
5. 正相关：高估斜率、低估截距，负相关反之。$E(\hat{\beta}_1) = E(\frac{x_i y_i}{\sum x_i^2}) = E(\beta_1 + \frac{\sum x_i \mu_i}{\sum x_i^2}) = \beta_1 + E(\sum \frac{x_i \mu_i}{\sum x_i^2}) \neq \beta_1$ ↩
6. 第一步：找到一个工具变量Z_2，对X_i估计，$X_i = \alpha_0 + \alpha_1 Z_{i1} + \alpha_2 Z_{i2} + \nu_i$（和2SLS第一步相同）；第二步：将上一步的残差加入原模型，OLS估计$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_{i1} + \delta \hat{\nu}_i + \varepsilon_i$ ↩
7. 第一步：$\tilde{\mu}_i = Y_i - (\tilde{\beta}_0 + \tilde{\beta}_1 X_i + \tilde{\beta}_2 Z_i)$；第二步：$\tilde{\mu}_i = \delta_0 + \delta_1 Z_{i1} + \delta_2 Z_{i2} + \delta_3 Z_i + \varepsilon_i$；统计量：$J = nR^2 \sim \chi^2(1)$（1表示多余的工具变量个数） ↩
8. 正规方程组：$\sum_{i=1}^n z_i \mu_i = 0$；估计量：$\tilde{\beta}_1 = \frac{\sum z_i y_i}{\sum z_i x_i}$。大样本下一致，小样本有偏 ↩
9. 原模型：$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \mu_i$；第一阶段：$\hat{X}_i = \hat{\alpha}_0 + \hat{\alpha}_1 Z_{i1} + \hat{\alpha} Z_{i2} + \hat{\alpha}_3 Z_i$，第二阶段：$Y_i = \hat{\beta}_0 + \hat{\beta}_1 \hat{X}_i + \hat{\beta}_2 Z_i + \mu_i$ ↩
10. 以$\tilde{e}_t$为被解释变量，选择各种可能的相关量（如$\tilde{e}_{t-1}$、$\tilde{e}_{t-2}$、$\tilde{e}_t^2$等）为解释变量，建立各种可能的回归方程，进行F检验 ↩
11. 原模型：$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \cdots + \beta_k X_{tk} + \mu_t$；零假设：$\mu_t = \rho \mu_{t-1} + \varepsilon_t$中$\rho = 0$；统计量：$D.W. = \frac{\sum_{t=2}^n (\tilde{e}_t - \tilde{e}_{t-1})^2}{\sum_{t=1}^n \tilde{e}_t^2} \approx 2(1-\rho)$ ↩
12. 辅助模型：$Y_t = \beta_0 + \beta_1 X_{t1} + \cdots + \beta_k X_{tk} + \rho_1 \mu_{t-1} + \cdots + \rho_p \mu_{t-p} + \varepsilon_t$；零假设：$\rho_1 = \cdots = \rho_p = 0$；统计量：$LM = nR^2 \sim \chi^2(p)$ ↩
13. $\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1} Y$ ↩

14. 原模型：$Y_t=\beta_0+\beta_1X_t+\mu_t==$，记为(1)式；假设存在一阶自相关$\mu_t=\rho\mu_{t-1}+\varepsilon_t$。将原模型滞后一期，$Y_{t-1}=\beta_0+\beta_1X_{t-1}+\mu_{t-1}$，记为(2)式，$(1)-\rho\times(2)$得$Y_t-\rho Y_{t-1}=\beta_0(1-\rho)+\beta_1(X_t-\rho X_{t-1})+(\mu_t-\rho\mu_{t-1})$。估计$\Delta Y_t=\beta_0'+\beta_1\Delta X_t+\varepsilon_t$。因为差分过程中，损失了第一个观测值，作普莱斯-温斯特变换$Y_1^*=\sqrt{1-\rho^2}Y_1, X_{1j}^*=\sqrt{1-\rho^2}X_{1j}$。代码：`equation eq02.ls y c x ar(1)`