动态面板回归

——基于GMM估计方法

报告人：冷学琪

武汉大学经济与管理学院

2019年10月31日

主要内容

- 动态面板回归模型
- GMM估计方法
  - 差分GMM
  - 水平GMM
  - 系统GMM
- 示例（Stata演示）
- GMM估计原理

1. 动态面板回归模型

1.1 动态面板基本形式

- 固定效应和双重差分估计假设

$$E(Y_{0it}|\alpha_i, X_{it}, D_{it}) = E(Y_{0it}|\alpha_i, X_{it})$$

重要的遗漏变量不随时间变化

- 动态面板回归模型基本形式

$$E(Y_{0it}|Y_{it-h}, \alpha_i, X_{it}, D_{it}) = E(Y_{0it}|Y_{it-h}, \alpha_i, X_{it})$$

$$Y_{it} = \alpha_i + \mu_t + \theta Y_{it-h} + \delta D_{it} + X'_{it}\beta + \varepsilon_{it}$$

其中 α_i是个体固定效应，μ_t是时间固定效应，Y_{it-h} 表示滞后多期的解释变量（向量），ε_{it}表示残差且与所有回归变量（Y_{it-1}和X_{it} ）无当期相关性。

1.2 动态面板的问题

- 差分消除固定效应

$$\Delta Y_{it} = \underbrace{\underline{\Delta Y_{it-1}} + \Delta \mu_t + \delta \Delta D_{it} + \Delta X_{it}{}'\beta + \underline{\Delta \varepsilon_{it}}}_{\text{相关性}}$$

解释：ΔY_{it-1}=$Y_{it-1} - Y_{it-2}$，$\Delta\varepsilon_{it}$=$\varepsilon_{it} - \varepsilon_{it-1}$，由于 Y_{it-1}与ε_{it-1}相关，$cov(\Delta Y_{it-1}, \Delta\varepsilon_{it}) \neq 0$

- 当面板回归具有动态特征时，静态面板的简单变换方法无法消除固定效应影响，因而OLS估计不可行，考虑引入相应的工具变量。

2. GMM估计方法

2.1 解决思路：滞后项作为工具变量

- 原理：Y_{it-2}作为ΔY_{it-1}的工具变量，然后进行2SLS估计(Anderson & Hsiao (1981))
- 假设：$\{\varepsilon_{it}\}$不存在自相关
- 解释： Y_{it-2}与ΔY_{it-1}相关，由于但与$\Delta \varepsilon_{it}$无关
- 问题： $\{Y_{it-3}, Y_{it-4}, \cdots\}$都是有效的工具变量，但是却未加利用，使得Anderson - Hsiao 估计量不是最有效率的估计量。

2.2 差分GMM

- 原理：所有滞后变量作为工具变量，采用GMM估计 (Arellano & Bond (1991))
  - 由于这一方法的基础是对差分方程进行估计，故称为差分GMM，文献中也大量引用“AB方法”这一名称。
- 假设：$\{\varepsilon_{it}\}$不存在自相关
- 优点：大幅度提高估计效率

2.3 差分GMM 的问题

- 弱工具变量问题
  - 滞后越多期，Y_{it-n}与ΔY_{it-1}相关性越弱。
  - 若$\{Y_{it}\}$有较强持续性（一阶自回归系数接近于1），Y_{it-2}与$\Delta Y_{it-1}= Y_{it-1} - Y_{it-2}$相关性较弱。
- 过度识别问题
- 差分带来的问题
  - 纯粹的截面作用α_i无法估计。
  - 降低信噪比 → 水平值时间、截面差异的信息都被弱化。

2.4 水平GMM

- 原理：$\{\Delta Y_{it-1}, \Delta Y_{it-2}, \cdots\}$作为工具变量对水平方程进行GMM 估计。(Arrelano & Bover (1995))
- 水平方程：

$$Y_{it} = \alpha_i + \mu_t + \theta Y_{it-h} + \delta D_{it} + X_{it}{}'\beta + \varepsilon_{it}$$

- 假设：$\{\varepsilon_{it}\}$不存在自相关；$\{\Delta Y_{it-1}, \Delta Y_{it-2}, \cdots\}$与个体效应$\alpha_i$不相关
- 解释：$\{\Delta Y_{it-1}, \Delta Y_{it-2}, \cdots\}$与$Y_{it-1}$相关；如果$\{\varepsilon_{it}\}$不存在自相关，则 $\mathrm{E}(\Delta Y_{it-s} \cdot \varepsilon_{it})$=0。

$$\mathrm{E}(\Delta Y_{it-s} \cdot \varepsilon_{it}) = \mathrm{E}(Y_{it-s} \cdot \varepsilon_{it}) - \mathrm{E}(Y_{it-s-1} \cdot \varepsilon_{it}) = 0$$

2.5 系统GMM

- 原理：结合差分GMM和水平GMM，将差分方程和水平方程作为一个方程系统，进行GMM估计。
- 假设：与水平GMM 相同
- 优点：
  - 提高估计效率
  - 估计不随时间变化的变量α_i的系数
  - 当样本自相关性比较高时，系统GMM的有限样本偏差要比差分GMM好。
- 缺点：当$\{\Delta Y_{it-1}, \Delta Y_{it-2}, \cdots\}$与个体效应$\alpha_i$相关时，无法使用这一模型。

2.6 模型设定检验

- 检验残差项ε_{it}的自相关性
  - 假设要求ε_{it}无自相关，所以$\Delta\varepsilon_{it}$不会具有2阶以上自相关
- 过度识别检验 → Sargan检验
  - 原假设为不存在过度识别
  - 当拒绝原假设时（p-值较小），需要考虑缩减工具变量的数量；或者是增多被解释变量滞后期的设定。

2.7 解释变量设定

- 前面假设的是解释变量X_{it} 完全外生：与ε_{it}及其滞后项均无相关性。
- 假设放松：允许存在前定变量（predetermined variable）和内生变量（endogenous variable）
- 前定变量: ω_{it}与ε_{it}不相关，但是与ε_{it-1}及更高阶滞后相关。
- 内生变量： ω_{it}与ε_{it}相关及更高阶滞后相关
- Question?

如何判断一个经济模型中的前定变量，内生变量和外生变量?

3. 示例演示

用STATA实现

3.1 数据来源及变量选择

- 数据：来自于Panel Study of Income Dynamics（PSID），包含了595名美国工人1976-1982年有关工资（n=595，T=7的短面板）
- 被解释变量：lwage（工资对数）
- 解释变量：
  - 外生变量：occ（是否是蓝领工人） south（是否在南方） smsa（是否在大城市） ind（是否在制造业）
  - 前定变量：wks（weeks worked，已工作周数）
  - 内生变量：ms（婚否）union（是否由工会合同确定工资）
  - 不随时间变化的变量：ed（教育年限）fem（是否女性） blk（是否是黑人）由于对原有模型进行了差分处理，这些变量将无法估计，故不包括在模型中。

3.2 模型设定

- 动态面板模型：

$$lwage_{it} = \alpha + \underbrace{\rho_1 \underline{lwage_{it-1}} + \rho_2 \underline{lwage_{it-2}}}_{\text{被解释变量滞后项}} +$$

$$\underbrace{\beta_1 \underline{occ}_{it} + \beta_2 \underline{south}_{it} + \beta_3 \underline{smsa}_{it} + \beta_4 \underline{ind}_{it}}_{\text{外生变量}} +$$

$$\underbrace{\beta_5 \underline{wks_{it}} + \beta_6 \underline{wks_{it-1}}}_{\text{前定变量}} +$$

$$\underbrace{\beta_7 \underline{ms_{it}} + \beta_8 \underline{union_{it}}}_{\text{内生变量}} + u_i + \varepsilon_{it}$$

4. GMM估计原理

4.1 GMM使用情景

- 球形扰动项假定下，2SLS是最有效率的。扰动项存在异方差或自相关，广义矩估计（Generalized Method of Moments，简记GMM）更有效。
- 总体矩条件

$$E(g_i) = E(z_i \varepsilon_i) = 0$$

- 样本矩条件

$$g_n(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^{n} z_i(y_i - x_i{}'\hat{\beta}) = 0$$

未知数$\hat{\beta}$共有K个，方程个数有L个（z_i的维度）

L<K，不可识别；L=K，恰好识别，有唯一解$\hat{\beta}_{IV}$；

L>K，过度识别，$\hat{\beta}$无解，传统的矩估计法无法求出。

4.2 GMM估计原理

- 利用“权重矩阵”W来构造二次型
  - $\widehat{W}$是$L \times L$维对称正定矩阵，$\operatorname*{plim}_{n\to\infty} \widehat{W} = W$，$W$为非随机的对称正定矩阵。
- 最小化目标函数

$$\min_{\hat{\beta}} J\left(\hat{\beta}, \widehat{W}\right) = n\left(g_n\left(\hat{\beta}\right)\right)' \widehat{W} \left(g_n\left(\hat{\beta}\right)\right)$$

$$\hat{\beta}_{GMM}\left(\widehat{W}\right) = \operatorname*{argmin}_{\hat{\beta}} J\left(\hat{\beta}, \widehat{W}\right)$$

- 最小化问题的解（求解过程略）

$$\hat{\beta}_{GMM}\left(\widehat{W}\right) = \left(S'_{ZX} \widehat{W} S_{ZX}\right)^{-1} S'_{ZX} \widehat{W} S_{Zy}$$

其中，$S_{ZX} = \frac{1}{n}\sum_{i=1}^{n} z_i x'_i$，$S_{Zy} = \frac{1}{n}\sum_{i=1}^{n} z_i y_i$

- 恰好识别下，S_{ZX}为方阵，GMM还原为工具变量法。

4.3 两步最优GMM估计和迭代法

- 最优GMM

  $\widehat{W}=\hat{S}^{-1}$是使$Avar(\hat{\beta}_{GMM})$最小化的最优权重矩阵，

  使用$\hat{S}^{-1}$为权重矩阵的GMM估计量为“最优GMM”

- 估计步骤

  - 使用2SLS，得到残差，计算$\hat{S}=\frac{1}{n}\sum_{i=1}^{n}(e_i)^2\, z_i z'_i$
  - 最小化$J(\hat{\beta},\hat{S}^{-1})$，得到$\hat{\beta}_{GMM}(\hat{S}^{-1})$

- 迭代法（实操中常用）

  用第二步所得到的残差再来计算$\hat{S}$，然后再来计算

  $\hat{\beta}_{GMM}(\hat{S}^{-1})$，直到估计值收敛。