动态面板回归的 GMM 方法

报告人：刘岩

武汉大学经管学院金融系

2017 年 6 月 6 日

概要

- 动态面板模型：在面板数据中考虑被解释变量的动态特征。
- 由于被解释变量的滞后项也进入回归方程，个体固定效应会导致普通的 OLS 回归产生偏误和不一致性。
  ——这也是回归内生性问题的一种形式。
- 为了克服 OLS 估计的问题，需要引入工具变量：在动态面板模型中，最常用的工具变量是被解释变量和解释变量的滞后及差分滞后项。
- 引入这类工具变量后，可利用 GMM 的一般框架进行估计，因此这类方法统称为动态面板回归的 GMM 方法。

本讲内容

1. 动态面板回归模型
2. GMM 估计方法
   - 差分 GMM
   - 系统 GMM
   - GMM 估计的选项与设定
3. 示例

本节内容

1. 动态面板回归模型
2. GMM 估计方法
   - 差分 GMM
   - 系统 GMM
   - GMM 估计的选项与设定
3. 示例

动态面板模型的基本形式

- 面板数据：截面 $i = 1, \dots, N$，时间 $t = 1, \dots, T$，样本 $\{y_{it}, \boldsymbol{x}_{it}\}$，$y_{it}$ 是被解释变量，$\boldsymbol{x}_{it}$ 是解释变量（向量）。
- 动态面板回归（dynamic panel regression）模型的最基本形式：

$$y_{it} = \phi y_{i t-1} + \boldsymbol{x}_{it}' \boldsymbol{\beta} + u_i + \varepsilon_{it},$$

  其中 $\boldsymbol{\beta}$ 是系数（向量），u_i 表示个体固定效应，ε_{it} 表示残差项且与所有回归变量（$y_{i t-1}$ 与 $\boldsymbol{x}_{it}$）无当期相关性。
  - 下面会进一步放松 $\boldsymbol{x}_{it}$ 与 ε_{it} 的相关性假设。
- 回归方程右侧可加入 y_{it} 的更多滞后项。

个体固定效应与内生性

- 在面板模型中，个体固定效应（individual fixed effect）u_i 用来捕捉无法观测的个体特征对被解释变量的影响，而这类个体异质性自然与个体观测变量 $\boldsymbol{x}_{it}$ 相关，$\mathrm{cov}(\boldsymbol{x}_{it}, u_i) \neq 0$。
  - 另一类捕捉个体异质性的方法，称为随机效应（random effect），其本质假设是 u_i 作用于 y_{it}，但 $\mathrm{cov}(\boldsymbol{x}_{it}, u_i) = 0$，即个体效应与可观测个体特征不相关。
  - 随机效应是一个非常强的假设，大部分严谨的实证研究都不会采用随机效应模型作为主要实证模型。
- 个体固定效应会导致动态面板回归产生内生性问题：回归变量与残差项 $u_i + \varepsilon_{it}$ 相关性不为 0。

固定效应：静态面板的情形

- 先考虑静态面板回归模型：

$$y_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} + u_i + \varepsilon_{it}.$$

- 此种情况下，为避免 u_i 引起的问题，可以通过简单的差分方法去除固定效应：

$$\Delta y_{it} = \Delta \boldsymbol{x}_{it}'\boldsymbol{\beta} + \Delta \varepsilon_{it},$$

  如此可以使用 OLS 进行估计。

- 也可以使用组内均值（within group）方法去除固定效应：令 $\bar{y}_i = \frac{1}{T}\sum_t y_{it}$，$\bar{\boldsymbol{x}}_i = \frac{1}{T}\sum_t \boldsymbol{x}_{it}$，$\bar{\varepsilon}_i = \frac{1}{T}\sum_t \varepsilon_{it}$，则有

$$y_{it} - \bar{y}_i = (\boldsymbol{x}_{it} - \bar{\boldsymbol{x}}_i)'\boldsymbol{\beta} + \varepsilon_{it} - \bar{\varepsilon}_i,$$

  亦可使用 OLS 估计。

固定效应：动态面板的问题

- 当面板回归模型具有动态特征时，静态面板的简单变换方法无法消除固定效应的影响，因而 OLS 估计不可行。
- 以差分方法为例：回归方程两边取差分后得

$$\Delta y_{it} = \phi \Delta y_{it-1} + \Delta \boldsymbol{x}'_{it} \boldsymbol{\beta} + \Delta \varepsilon_{it},$$

  但 $\mathrm{cov}(\Delta y_{it-1}, \Delta \varepsilon_{it}) \neq 0$，$\varepsilon_{it-1}$ 对 y_{it-1} 有影响。
- 组内平均的变换带来的问题更大。

本节内容

1 动态面板回归模型

2 GMM 估计方法

- 差分 GMM
- 系统 GMM
- GMM 估计的选项与设定

3 示例

基本解决思路：滞后项作为工具变量

- 差分法在动态面板下无法克服固定效应带来的问题：根源在于被解释变量的动态性 $\Rightarrow t$ 期回归变量 y_{it-1} 与残差差分项中 ε_{it-1} 的相关性。
- 如果能找到一个变量，与差分方程中的回归变量相关，但与残差无关，则可用作工具变量。
- 最简单的工具变量：y_{it-2}，其与 Δy_{it-1} 自然相关，但与 $\Delta \varepsilon_{it}$ 无关！
  - 这一方法最早由 Anderson & Hsiao (1981) 提出。
- **基本假设 1**：ε_{it} 没有序列相关性。

用尽滞后期：差分 GMM

- Arrelano & Bond (1991, RES) 提出把所有的滞后项全部引入为工具变量，使用 GMM 方法进行估计。
  - 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；同时文献中也大量使用 AB 方法这一名称。
- 对被解释对 t 期回归方程，可用的工具变量有 $y_{it-2}, \ldots, y_{i1}$；滞后期工具变量总数可达 $T^2/2$ 量级 $\Rightarrow$ 可能引起过度识别（over-identification）问题。
- 如此一来，有可能大幅提高估计的效率。
  - Arrelano & Bond (1991) 是引用率最高的几篇计量文章之一，Google Scholar 引用超 20,000；可能仅低于 Engle & Granger (1987) 协整序列的文章。

差分 GMM 的问题

1. 因为取了差分，所以纯粹截面作用无法估计。
2. 差分可能会降低信噪比 $\Rightarrow$ 水平值时间、截面差异的信息都被弱化。
3. 如果 y_{it} 和解释变量的一阶自相关很接近 1，则可能存在弱工具变量问题。
   - 若 $\phi \approx 1$，那么 $\Delta y_{it} = y_{it} - y_{it-1}$ 几乎完全由回归方程中的余项决定，与 y_{it-1} 的相关系数会很小。

水平回归与差分滞后工具变量

- 针对上述问题，特别是高自相关性带来的弱工具变量问题 3，Arrelano & Bover (1995) 提出了另外一种选取工具变量的方法：水平回归与差分滞后。
- 回到动态面板水平（level）回归方程：

$$y_{it} = \phi y_{it-1} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + u_i + \varepsilon_{it}.$$

如果被解释变量差分滞后项 Δy_{it-1} 与固定效应 u_i 没有相关性，那么 Δy_{it-1} 也可以作为工具变量！

系统 GMM

- **基本假设 2**：$\text{cov}(\Delta y_{it-1}, u_i) = 0$。
- 在此假设下，差分回归与滞后工具变量可以同水平回归与差分滞后工具变量相结合，形成一个更完善的 GMM 估计系统。这一方法被 Blundell & Bond (1998) 称为系统 GMM。
  - Blundell & Bond 给出的模拟结果显示，当样本自相关很高时，系统 GMM 的有限样本偏差要比差分 GMM 好。
- Blundell & Bond 同时指出了基本假设 2 的实质：给定个体效应 u_i，y_{it} 的长期均值主要由 $u_i/(1-\phi)$ 决定；在此基础上，基本假设 2 等价于

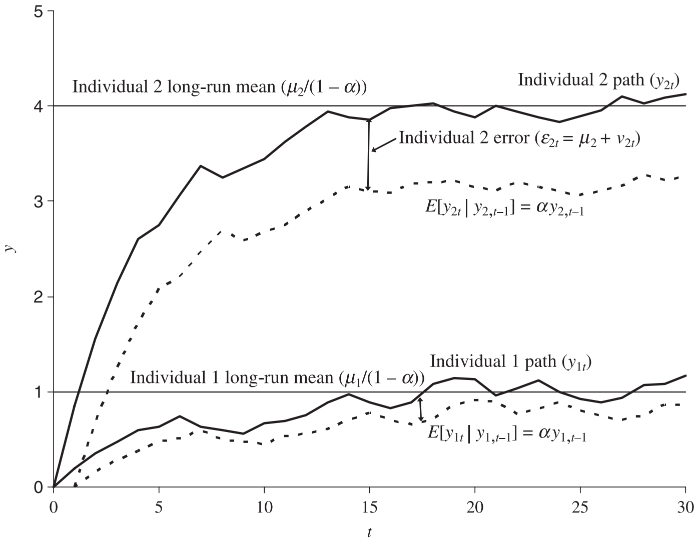
$$\mathbb{E}\left[\left(y_{it} - \frac{u_i}{1-\phi}\right)u_i\right] = 0, \qquad (*)$$

对 t 成立，即 y_{it} 关于其长期均值的暂时偏离与 u_i 无关。

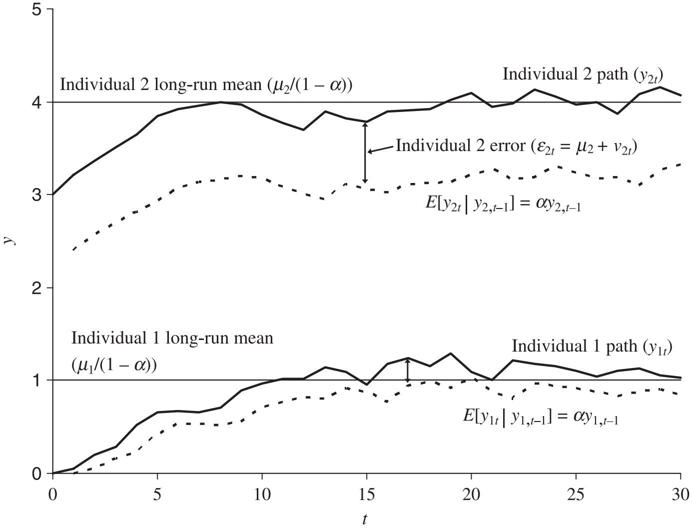
基本假设的进一步解释

- 由 y_{it} 的自回归性质可知，条件 $(*)$ 只需对初始样本 y_{i1} 成立即可。
  - 故 Blundell & Bond 将该条件称为初值条件。
- 进一步的，该条件可以理解为当面板样本初始截面与个体长期均值的偏离 $y_{it} - \frac{u_i}{1-\phi}$，平均而言与各自长期均值的相关性为 0；换言之，个体的初始偏离水平与长期趋势无关。
- 下面两幅图说明对此进行了说明；来源：Roodman (2009)。

不满足差分工具变量的情形



满足差分工具变量的情形



GMM 估计的基本方法：以系统 GMM 为例

- 定义样本矩阵：$*$ 表示样本变换（如差分），L 表示水平值，

$$\boldsymbol{y}_i = \begin{bmatrix} \boldsymbol{y}_i^* \\ \boldsymbol{y}_i^L \end{bmatrix} \qquad \boldsymbol{X}_i = \begin{bmatrix} \boldsymbol{X}_i^* \\ \boldsymbol{X}_i^L \end{bmatrix},$$

$\boldsymbol{Z}_i$ 表示所有工具变量构成的矩阵。

- 给定变换矩阵 $\boldsymbol{H}_i$，定义

$$\boldsymbol{Q}_{xz} = \sum_i \boldsymbol{X}_i' \boldsymbol{Z}_i, \quad \boldsymbol{Q}_{zy} = \sum_i \boldsymbol{Z}_i' \boldsymbol{y}_i, \quad \boldsymbol{W} = \boldsymbol{Q}_{xz} \boldsymbol{A} \boldsymbol{Q}_{xz}',$$

$$\boldsymbol{A} = \left(\sum_i \boldsymbol{Z}_i' \boldsymbol{H}_i \boldsymbol{Z}_i \right)^{-1}.$$

- 回归系数的 GMM 估计值为

$$\hat{\boldsymbol{\theta}} = \boldsymbol{W}^{-1} \boldsymbol{Q}_{xz} \boldsymbol{A} \boldsymbol{Q}_{zy}.$$

一步法与两步法

- 动态面板 GMM 估计分为一步法与两步法。
- 在一步法中，变换矩阵 $\boldsymbol{H}_{1i}$ 取特定常数值，对应的加权矩阵（weighting matrix）$\boldsymbol{A}_1$ 直接计算可得。
- 在两步法中，变换矩阵 $\boldsymbol{H}_{2i}$ 由一步法估计值 $\hat{\boldsymbol{\theta}}_1$ 计算残差构造而得：

$$\hat{\boldsymbol{\varepsilon}}_i = \boldsymbol{y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\theta}}_1, \qquad \boldsymbol{H}_{2i} = \hat{\boldsymbol{\varepsilon}}_i' \hat{\boldsymbol{\varepsilon}}_i.$$

再由 $\boldsymbol{H}_{2i}$ 计算新的加权矩阵 $\boldsymbol{A}_2$，及相应的 $\boldsymbol{W}_2$，最终得到两步法估计值 $\hat{\boldsymbol{\theta}}_2 = \boldsymbol{W}_2^{-1} \boldsymbol{Q}_{xz} \boldsymbol{A}_2 \boldsymbol{Q}_{zy}$。

GMM 估计值的统计推断：一步法下 $\hat{\theta}$ 协方差矩阵

- 在同方差假设下，一步法得到的参数估计（向量）$\hat{\boldsymbol{\theta}}_1$ 具有下列协方差矩阵：$\hat{\sigma}_1^2 \boldsymbol{W}_1^{-1}$，其中 $\hat{\sigma}_1^2$ 表示残差（变换后）的同方差估计值。
- 一步法下，可以使用 White 的异方差稳健标准误（robust standard error）方法估计 $\hat{\boldsymbol{\theta}}_1$ 的协方差矩阵：

$$\boldsymbol{W}_1^{-1} \boldsymbol{Q}_{xz} \boldsymbol{A}_1 \boldsymbol{A}_2^{-1} \boldsymbol{A}_1 \boldsymbol{Q}'_{xz} \boldsymbol{W}_1^{-1},$$

其中 $\boldsymbol{A}_2$ 是两步法中使用的加权矩阵（由一步法残差构造）。

GMM 估计值的统计推断：两步法下 $\hat{\theta}$ 协方差矩阵

- 两步法下，同方差假设下 $\hat{\boldsymbol{\theta}}_2$ 具有协方差矩阵 $\boldsymbol{W}_2^{-1}$。
- 但由于一步法残差构造的两步法加权矩阵有很大的有限样本偏误，利用 $\boldsymbol{W}_2^{-1}$ 所做统计推断均存在很大偏差。
  - 150 左右的样本即可保证所有系数 *** 显著。
- Windermeijer (2005) 提出了经过偏误纠正的加权矩阵估计值 $\boldsymbol{A}_{2,\text{robust}}$，对应 $\hat{\boldsymbol{\theta}}_2$ 的协方差矩阵称为偏误纠正（或 WC-robust）协方差阵。
- WC-robust 协方差阵得到的统计推断结果具有较好的有限样本性质；综合水平与一步法稳健标准误结果类似或略高。

模型设定检验

- 对动态面板 GMM 的模型设定检验通常包括两类。
- 首先是对残差项 ε_{it} 自相关性的检验。基本假设 1 要求 ε_{it} 无自相关，故 $\Delta\varepsilon_{it}$ 不会具有 2 阶及以上自相关。
- 其次是对工具变量数目过多带来的过渡识别进行检验，通常可以进行一个 Sargan 检验：原假设为不存在过渡识别，因此当以较小的 p-值拒绝原假设时，需要考虑缩减工具变量的数量。
  - 但其他方面的模型设定问题也可能引起 Sargan 检验拒绝原假设，如被解释变量滞后期设定不足等。

解释变量设定

- 前面的讨论中我们假定解释变量 $\boldsymbol{x}_{it}$ 完全外生：与 ε_{it} 及其滞后项均无相关性。
- 上述假设可以放松：允许存在前定变量（predetermined variable）和内生变量（endogenous variable）。
- 前定变量：w_{it} 与 ε_{it} 不相关，但可以与 ε_{it-1} 及更高阶滞后相关。
- 内生变量：w_{it} 与 ε_{it} 相关及更高阶滞后相关。
- 在 GMM 估计设定中，可以指定哪些变量是前定变量或内生变量，软件会相应的引入这些变量的对应水平滞后或者差分滞后作为工具变量，进入 GMM 估计程序。

本节内容

1 动态面板回归模型

2 GMM 估计方法

差分 GMM

系统 GMM

GMM 估计的选项与设定

3 示例

金融对经济增长的影响

- Beck & Levine 2004 JBF, Arcand et al. 2015 JEG.
- 跨国面板回归：

$$y_{it} - y_{it-1} = \phi y_{it-1} + \beta F_{it} + \boldsymbol{x}'_{it}\boldsymbol{\theta} + \varepsilon_{it},$$

  其中 y_{it} 是对数人均 GDP，F_{it} 是金融发展指标，被解释变量是经济增长率。

- 这个回归自然具有动态面板的特征，使用 GMM 方法可以获得最有效的估计。