第5章 动态计量经济学模型

5.1从数据生成过程到自回归分布滞后模型

5.2从自回归分布滞后模型到误差修正模型

本章说明

- 70年代末80年初，以英国计量经济学家D.F.Hendry为代表，在误差修正模型的基础上，提出了动态计量经济学模型的理论与方法。
- 动态建模理论具有与传统的经典建模理论不同的思路。

- 经典建模理论的模型设定理论可以概括为：
  - 依据某种已经存在的经济理论或者已经提出的对经济行为规律的某种解释设定模型的总体结构和个体结构，即模型是建立在已有的经济理论和经济行为规律假设的基础之上的；
  - 引进概率论思想作为模型研究的方法论基础，选择随机联立线性方程组作为模型的一般形式；
  - 模型的识别、参数的估计、模型的检验是主要的技术问题；
  - 以模型对样本数据的拟合优度作为检验模型的主要标准。

- 经典建模理论的模型设定方法的特点为：
  - 从简单到复杂
  - 从一般到简单
  - 所谓“一般”不同于动态模型中的“一般”

- Lucas批判
  - Lucas(1976)、Sarget(1976)、Sims(1980)
  - Lucas(1976)：使用计量经济模型预测未来经济政策的变化所产生的效用是不可信的。其实质是提出了结构模型模型参数是否随时间变化的问题。
  - Sarget(1976)：以货币政策为例，重新解析了Lucas批判。结构模型对于评价政策似乎是无能为力的。
  - Sims(1980)：为使结构方程可以识别而施加了许多约束，这些约束是不可信的。建议采用向量自回归（VAR）模型而避免结构约束问题。
  - 关于模型设定：经济学理论不足以指导如何设定模型，以及保证模型设定的正确性。

- 背景
  - 20世纪70年代的世界经济
    - 滞涨
    - 石油危机
    - 利率自由化
    - 管理浮动汇率
  - 关于宏观经济政策有效性的争论
    - 以弗里德曼为代表货币主义的固定规则
    - 以卢卡斯、萨金特、华莱士等为代表新古典宏观经济学第一代的货币政策无效
    - 以基德兰德、普利斯科特等为代表新古典宏观经济学第二代的财政政策无效
    - 新凯恩斯主义的货币政策连续性

- 动态计量经济学—Hendry学派建模理论：
  - 数据、理论双导向——交替运用经济理论和经济数据提供的信息。
  - 首先建立一个能够代表数据生成过程(DGP, Data Generation Process)的自回归分布滞后模型(ADL, Autoregressive Distributed Lag)，然后逐步简化，最后得到包含变量间长期稳定关系的简单的(ECM, Error Correct Model)模型。

- 从DGP到ADL：

$$\prod_{t=1}^{T} D(x_t \mid x_{t-1};\theta)$$

数据生成过程(DGP)是所有变量联合概率分布的一般表达式。

每步约化都有效

通常所说的“从一般到简单”的“一般”，是建立模型的起点。

$$y_t = \beta_0 + \sum_{i=0}^{q} \delta_i z_{t-i} + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t$$

- 从ADL到ECM：

$$y_t = \beta_0 + \sum_{i=0}^{q} \delta_i z_{t-i} + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t$$

每步简化都有效

$$\Delta y_t = \tilde{\beta}_0 + \sum_{i=0}^{q} \tilde{\delta}_i \Delta z_{t-i} + \sum_{i=1}^{p} \tilde{\gamma}_i \Delta y_{t-i} + \alpha\, ecm_{t-1} + \varepsilon_t$$

- 将本章主要内容分为3讲：
  - 从DGP到ADL
  - ADL
  - 从ADL到ECM

已经放到本科教学内容中，不讲。

5.1 从数据生成过程到自回归分布滞后模型

一、数据生成过程(DGP)

二、弱外生性、强外生性和超外生性

三、约化

四、自回归分布滞后模型和数据生成过程

- 动态计量经济模型的起点是反映数据生成过程（**DGP**）的自回归分布滞后模型（**ADL**），那么，从**DGP**到**ADL**，成为动态计量经济模型理论方法的一个重要部分。
- 只涉及思路，不进行数学描述。

一、数据生成过程(DGP)

数据生成过程的表达、含义、分解

- 数据生成过程，是描述已经得到的变量观测值是如何生成的。
- 可以用所有变量（包括内生变量和外生变量）观测值的联合概率密度函数来表示。

设 x_t 是 t 时期所有变量观测值向量，用 $\mathbf{X}_{t-1}$ 表示从第 1 时期到第 $t-1$ 时期所有变量观测值矩阵，即 $\mathbf{X}_{t-1}=(x_{t-1},x_{t-2},\cdots,x_1)'$，那么观测值 x_t 的联合概率密度函数，即 DGP，可以表示为：

$$\prod_{t=1}^{T} D(x_t | \mathbf{X}_{t-1}; \Theta)$$

X_t的发展变化只依赖于自身的历史X_{t-1}，而与其它因素无关。

如果x_t可以分为内生变量y_t和外生变量z_t， $\mathbf{X}_{t-1}$也作如下相应分解：

$$\mathbf{X}_{t-1} = (\mathbf{Y}_{t-1} : \mathbf{Z}_{t-1})$$

$$\mathbf{Y}_{t-1} = (y_{t-1}, y_{t-2}, \cdots, y_1)'$$

$$\mathbf{Z}_{t-1} = (z_{t-1}, z_{t-2}, \cdots, z_1)'$$

联合概率密度函数可以分解为 y_t 对 z_t 的**条件密度** $D_{y|z}$

与 z_t 的**边际密度** D_z 的乘积：

$$D(x_t|\mathbf{X}_{t-1};\Theta)=D_{y|z}(y_t|z_t,\mathbf{Y}_{t-1},\mathbf{Z}_{t-1};\lambda_1)\cdot D_z(z_t|\mathbf{Y}_{t-1},\mathbf{Z}_{t-1};\lambda_2)$$

其中 λ_1 和 λ_2 分别为条件密度和边际密度的参数向量，是由是联合概率密度函数 D 的原未知参数向量 Θ 分解而来。如果 Θ 中包括 n 个参数，λ_1 和 λ_2 中分别包括 n_1 和 n_2 个参数，则应有 $n = n_1 + n_2$。

二、弱外生性、强外生性和超外生性

1. 弱外生性的条件

- 如果能够将外生变量z_t的边际密度D_Z合理地约去，则可以使得描述数据生成过程的联合概率密度函数简化。其条件是z_t外生变量具有弱外生性。
- 所有参数集Θ，关注的一个子集Ψ，在Ψ中包括的参数数目$k \leqslant n_1$，这些参数描述了内生变量y_t与相关联的变量之间的关系。子集Ψ被称为关注参数。

$$D(x_t | \mathbf{X}_{t-1}; \Theta) = D_{y|z}(y_t | z_t, \mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}; \lambda_1) \cdot D_z(z_t | \mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}; \lambda_2)$$

- 如果能够将外生变量z_t的边际密度D_t从联合概率密度函数 D中合理地约去，那么Ψ应该仅为λ $_1$的函数而与λ $_2$无关。这就构成具有弱外生性的第1个条件。

$$\Psi = g(\lambda_1)$$

$$D(x_t | \mathbf{X}_{t-1};\Theta) = D_{y|z}(y_t | z_t, \mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}; \lambda_1) \cdot D_z(z_t | \mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}; \lambda_2)$$

- 在满足第1条件的情况下，如果λ_1对λ_2存在某种依存性，那么仍然不能将外生变量z_t的边际密度D_t从联合概率密度函数 D中合理地约去。于是引出了z_t具有弱外生性的第2个条件，即λ_1与λ_2相互无关。

该条件实际上是对参数λ_1与λ_2的子空间Λ_1与λ_2的某种约束，要求该两组参数之子空间的积满足：

$$\Lambda_1 \times \Lambda_2 = \{(\lambda_1, \lambda_2) : \lambda_1 \in \Lambda_1 \text{和} \lambda_2 \in \Lambda_2\}$$

- 由于第1和第2条件断绝了关注参数Ψ直接或者间接依存于λ_2的可能性，因此保证了关于z_t的边际密度D_t的信息对于估计关注参数Ψ的无关性。
- 所以z_t的弱外生性是从联合概率密度函数 D中合理地约去z_t的边际密度D_t的充要条件。
- 弱外生性的定义：如果1条件和第2条件成立，即z_t为关注参数Ψ的弱外生变量。该定义表明，变量的弱外生性必须是相对于给定的关注参数而言的。
- 弱外生性不是变量本身的性质 。

注意与经典模型的区别

- 例题

对于一个简单的粮食供求模型：

$$\begin{cases} p_t = \alpha_0 + \alpha_1 q_t + \mu_t & \mu_t \sim N(0, \sigma_\mu^2) \\ q_t = \beta_0 + \beta_1 p_{t-1} + \varepsilon_t & \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \end{cases}$$

$$E(\mu_t \varepsilon_t) = 0 \qquad \forall t, s$$

需求方程参数集为：$\lambda_1 = (\alpha_0, \alpha_1, \sigma_\mu^2)$；供给方程的参数集为：$\lambda_2 = (\beta_0, \beta_1, \sigma_\varepsilon^2)$。$q_t$ 是否是弱外生性变量？

如果关注参数为需求弹性,即价格与需求量之间的关系,可以用$1/\alpha_1$近似表示，显然只是λ_1的子集，q_t是弱外生性变量。

如果关注参数为系统的稳定性，即需要将供给方程代入需求方程，得到

$$p_t = \alpha_0 + \alpha_1\beta_0 + \alpha_1\beta_1 p_{t-1} + \nu_t$$

表示为：
$$p_t = \gamma + \rho\, p_{t-1} + \nu_t$$

由经济理论知道，参数 ρ 反映系统的稳定性，只有当 $|\rho|<1$ 时，模型系统收敛于一个均衡价格。显然此时的关注参数不仅需要 $\alpha_1 \in \{\lambda_1\}$，而且需要 $\beta_1 \in \{\lambda_2\}$，所以 q_t 不是弱外生性变量。

- 弱外生性不是变量本身的性质，是相对于关注参数而言的。

2. 强外生性的条件

- 弱外生性仅相对于利用模型进行关注参数的统计推断的情形，强外生性是相对于利用模型进行预测的情形。
- 如果要利用联合概率密度函数进行预测并能够合理地约去z_t的边际密度D_t，仅有z_t的弱外生性是不够的，还必须限定z_t不受Y_{t-1}的反馈影响，这就是z_t的强外生性。

- 强外生性定义：

在z_t为关注参数Ψ的弱外生变量的基础上，如果下述条件成立，则z_t为关注参数Ψ的强外生变量。

$$D_z(z_t \mid \mathbf{X}_{t-1};\lambda_2) = D_z(z_t \mid \mathbf{Z}_{t-1}^1, \mathbf{X}_0;\lambda_2)$$

3. 超外生性的条件

- 超外生性是相对于利用模型进行政策分析的情形。
- 超外生性条件为：在z_t为关注参数Ψ的弱外生变量的基础上，如果λ_1对于属于子集C^2的政策干预具有抗变性，则z_t为关注参数Ψ的超外生变量。
- 超外生性是以弱外生性为基础，并没有以强外生性为基础。即是说超外生性包含弱外生性，但不包含强外生性。

C^2是能够引起来Θ变动的干预集C^{Θ}的一个子集，是仅对λ_2产生影响的一类可行干预。

- 从直观上看，如果条件模型中的解释变量具有强外生性，那么被控制变量（内生变量）的滞后项对它没有影响。
- 然而，许多政策干预是政府根据被控制变量的滞后信息而制定的。这说明，具有强外生性的解释变量一般不应被选择为实施政策干预的超外生变量。
- 重要提示：预测模型和政策评价模型是两类模型，同一个模型不可能同时具有两方面功能。

三、约化

- 约化理论是动态计量经济学模型理论的重要组成部分。由联合概率密度函数表示的数据生成过程经过约化，得到作为应用计量经济学模型起点的自回归分布滞后模型，只有在约化过程中没有关于关注参数的信息损失，自回归分布滞后模型才能代表数据生成过程。
- 但是，在实际建模中，人们往往直接从自回归分布滞后模型开始，并不实际地进行数据生成过程的约化。

1. 条件化—分布的约化

- 使得描述数据生成过程的联合概率密度函数简化为$\mathbf{y}_t$对$\mathbf{z}_t$的条件密度的过程。
- 关键是
  - 将$\mathbf{x}_t$分为内生变量$\mathbf{y}_t$和外生变量$\mathbf{z}_t$；
  - 将联合概率密度函数表示为$\mathbf{y}_t$对$\mathbf{z}_t$的条件密度与$\mathbf{z}_t$的边际密度的乘积；
  - $\mathbf{z}_t$为关注参数Ψ的弱外生变量。
- 数据生成过程的联合概率密度函数简化为：

$$D_{y|z}(y_t|z_t, \mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}; \lambda_1)$$

2. 新生化—误差项的约化

$$u_t = E(y_t \mid z_t, \mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}; \lambda_1)$$

$$\varepsilon_t = y_t - u_t$$

$$E(\varepsilon_t \mid \mathbf{Z}_t, \mathbf{Y}_{t-1}) = E(y_t \mid \mathbf{Z}_t, \mathbf{Y}_{t-1}) - u_t = 0$$

- 如果上述条件化是合理的，该新生误差应该是白躁声。

3. 常数化—参数的约化

- 在样本数据为时间序列数据时，需要对参数施加时间齐次性约束。这是数据生成过程向应用模型靠拢的关键一步。当模型的参数具有常数性时，有：

$$\lambda_{1t} = \lambda_1 \qquad \forall t; \quad \lambda_1 \in \Lambda_1$$

- 常数化过程是通过变量的选择来实现的。

- 变参数模型可以通过引入政策变量使之变为常参数模型。当然这里引入的变量应该具有超外生性。

4. 截尾化—滞后项的约化

$$D_{y|z}(y_t | z_t, \mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}; \lambda_1) = D_{y|z}(y_t | z_t, \mathbf{Y}_{t-1}^{t-p}, \mathbf{Z}_{t-1}^{t-q}; \lambda_1)$$

- 对条件密度中的滞后项进行滞后项截尾。
- 这一约化过程合理性的条件是新生误差是白躁声。

$$\boxed{\varepsilon_t = y_t - E(y_t | z_t, \mathbf{Y}_{t-1}^{t-p}, \mathbf{Z}_{t-1}^{t-q}; \lambda_1)}$$

5. 线性化—函数形式的约化

- 对变量y_t和z_t的某种函数转换，使得转换后的条件分布近似服从同方差正态分布。

$$\boxed{y_t^* = h(y_t) \qquad z_t^* = g(z_t)}$$

$$D_{y|z}(y_t|z_t, \mathbf{Y}_{t-1}^{t-p}, \mathbf{Z}_{t-1}^{t-q}; \lambda_1) = D_{y^*|z^*}(y_t^*|z_t^*, \mathbf{Y}_{t-1}^{*t-p}, \mathbf{Z}_{t-1}^{*t-q}; \gamma) \underset{app}{\sim} N(\eta_t, \sigma^2)$$

在上式成立的条件下，y_t^*基于z_t^*的条件模型便是线性的。

四、自回归分布滞后模型和数据生成过程

- 数据生成过程经过上述约化后，得到它的衍生模型，如果每一步约化都是有效的，即没有关于关注参数的信息损失，那么该衍生模型代表了数据生成过程。

- 利用滞后算子来表示衍生模型：

$$\mathbf{A}(L)\mathbf{y}_t^* = \mathbf{B}(L)\mathbf{z}_t^* + \varepsilon_t \qquad\qquad \varepsilon_t \underset{app}{\sim} N_{n1}(0, \Sigma_\varepsilon)$$

- 约化得到的衍生模型实际上是向量自回归分布滞后模型。误差项不是源生的，而是模型衍生的。

- 该模型就成为代表数据生成过程的动态计量经济学模型的起点。

讨论

- 在实际建模时，直接以自回归分布滞后模型为起点，并不讨论上述约化过程。
- 那么，如何才能保证所选择的 **ADL**模型满足上述约化的要求呢？关于线性化，采用经典模型的方法处理；关于截尾化，下节还要简单介绍；关于常数化，更多的是一种假设；关于新生化，主要依靠变量选择，尽可能将对被解释变量具有影响的变量选入模型；关于条件化，即变量关于参数的弱外生性，主要依靠对经济行为的理解加以确定。

5.2 从自回归分布滞后模型到误差修正模型

一、自回归分布滞后模型的阶数的决定

二、正交化变换

三、协整检验

四、协整与误差修正模型

一、自回归分布滞后模型的阶数的决定

1. 变量及滞后阶的选择

$$y_t = \beta_0 + \sum_{i=0}^{q} \delta_i z_{t-i} + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t$$

- 模型中变量的选择，仍然需要从经济理论和对经济行为规律的认识出发，人们不会也不应该将毫无关系的变量选入模型。正是在这点上，体现了动态计量经济学模型离不开理论导向。
- 理论上讲，决定模型中变量滞后的最大阶数的依据是误差项为一新生过程。

- 在实际中，人们更多的是从经验出发判断阶数。
  - 如果采用季节数据，滞后阶数至少大于**4**；如果采用年度数据，滞后阶数至少大于**1**。
  - 如果**y**表示消费，**z**表示收入，那么**(2,2)**较为合适，因为消费的调整比较容易；
  - 如果**y**表示投资，**z**表示收入，那么滞后阶应该比较高，因为投资的调整比较困难。
  - 在实践中，也可以开始设定较长的滞后阶数，然后对模型采用**OLS**回归，根据变量的显著性检验确定选择的阶数。

2、Akaike's FPE Criterion

- 利用**Akaike**最终预测差（**FPE, Final Prediction Error**）标准为基础确定最优滞后期。

$$y_t = \beta_0 + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t$$

给定不同的p值进行估计

$$FPE(p) = \frac{n+p+1}{n-p-1} \times ESS(p)/n$$

其中**ESS(p)**为不同的**p**值时的**OLS**估计残差平方和。选择使得**ESS(p)**最小的**p***为**y**的最优滞后期。

- 将y的最优滞后期p*作为给定值代入模型

$$y_t = \beta_0 + \sum_{i=0}^{q} \delta_i z_{t-i} + \sum_{i=1}^{p^*} \gamma_i y_{t-i} + \varepsilon_t$$

给定不同的q值进行估计

$$FPE(p^*, q) = \frac{n + p^* + 1}{n - p^* - 1} \times ESS(p^*, q)/n$$

选择使得**ESS(p*,q)**最小的**q***为**z**的最优滞后期。

二、正交化变换

- 在确定了模型中变量滞后的最大阶数后，对一般模型施加某种变换，使解释变量之间近似正交。
- 在已经见到的动态模型中，一般都将正交变换过程与引入误差修正机制结合起来，最后得到解释变量具有近似正交性的误差修正模型。
- 从理论上，下列过程可以实现正交变换。

- 下面以简单线性模型为例，说明正交变换过程。

设有线性模型： $y_t = \beta' \mathbf{Z}_t + \varepsilon_t$

设 $\mathbf{H}$ 为任意一个已知的 $(k \times k)$ 阶非奇异矩阵，满足 $\mathbf{H}^{-1}\mathbf{H} = \mathbf{I}$， γ 为一个已知的 $(k \times 1)$ 阶向量。首先从模型)两边减去 $\gamma' \mathbf{Z}_t$： 得到：

$$y_t - \gamma' \mathbf{Z}_t = (\beta - \gamma)' \mathbf{Z}_t + \varepsilon_t$$

然后，变换为：

$$(y_t - \gamma' \mathbf{Z}_t) = \{(\beta - \gamma)' \mathbf{H}^{-1}\} \mathbf{H}\mathbf{Z}_t + \varepsilon_t$$

即 $$y_t^* = \beta^{*'} \mathbf{Z}_t^* + \varepsilon_t$$

其中， $\beta^* = \mathbf{H}^{-1}(\beta - \gamma)$ ， $y_t^* = (y_t - \gamma' \mathbf{Z}_t)$， $\mathbf{Z}_t^* = \mathbf{H}\mathbf{Z}_t$， 实际上是对原始模型进行了再参数化，仍然有相同数目的参数、相同数目的回归元和相同的误差项。

至于是首先估计原始得到 $\hat{\beta}$ 然后导出 $\hat{\beta}^*$，还是直接估计变换后的模型得到 $\hat{\beta}^*$，是无所谓的。因为

$$
\begin{aligned}
\hat{\beta}^* &= (\mathbf{Z}^{*\prime}\mathbf{Z}^*)^{-1}\mathbf{Z}^{*\prime}y^* = (\mathbf{H}'\mathbf{Z}'\mathbf{Z}\mathbf{H})^{-1}\mathbf{H}'\mathbf{Z}'(y-\gamma\mathbf{Z}) \\
&= \mathbf{H}^{-1}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(y-\gamma\mathbf{Z}) = \mathbf{H}^{-1}(\hat{\beta}-\gamma)
\end{aligned}
$$

这就是正交变换的过程。

- 经过正交变换的模型的解释变量是近似正交的，该模型有一个显著的优点是，在模型简化的过程中，当去掉系数估计值小的变量或者去掉显著性差的变量时，在数值上和统计上，几乎不改变剩余变量的系数估计值。
- 所以首先对模型进行正交变换，然后再对模型进行简化，应该是一条正确的建模路线。
- 关键是选择合适的**H**矩阵。这并非容易的事情。

三、协整检验

（已经放在第1章）

四、协整与误差修正模型

1. 用误差修正模型表示协整体系

$$\Delta \mathbf{y}_t = \sum_{j=1}^{p} \Gamma_j \Delta \mathbf{y}_{t-j} + \alpha\beta' \mathbf{y}_{t-1} + \varepsilon_t$$

$$\downarrow$$

$$\Delta \mathbf{y}_t = \sum_{j=1}^{p} \Gamma_j \Delta \mathbf{y}_{t-j} + \theta\, \mathbf{ECM}_{t-1} + \varepsilon_t$$

- 其中每个方程都是一个误差修正型方程，可以剔除一些不显著的滞后差分项。误差修正项反映变量之间的关系偏离长期均衡状态对短期变化的影响，所有作为解释变量的差分项反映各变量短期变化对作为被解释变量的短期变化的影响。

2. 动态计量经济学模型—误差修正模型

$$\Delta y_t = \tilde{\beta}_0 + \sum_{i=0}^{q} \tilde{\delta}_i \Delta z_{t-i} + \sum_{i=1}^{p} \tilde{\gamma}_i \Delta y_{t-i} + \alpha\, ecm_{t-1} + \varepsilon_t$$

Hendry仅将它作为一种模型形式；**Granger**揭示了它的经济学和统计学内涵。