政治经济学前沿方法论与量化分析

第六讲 多元和非平稳时间序列的建模

上课地点：　善斋306C
上课时间：周二第六大节

龙治铭
善斋307C
zhiminglong@tsinghua.edu.cn

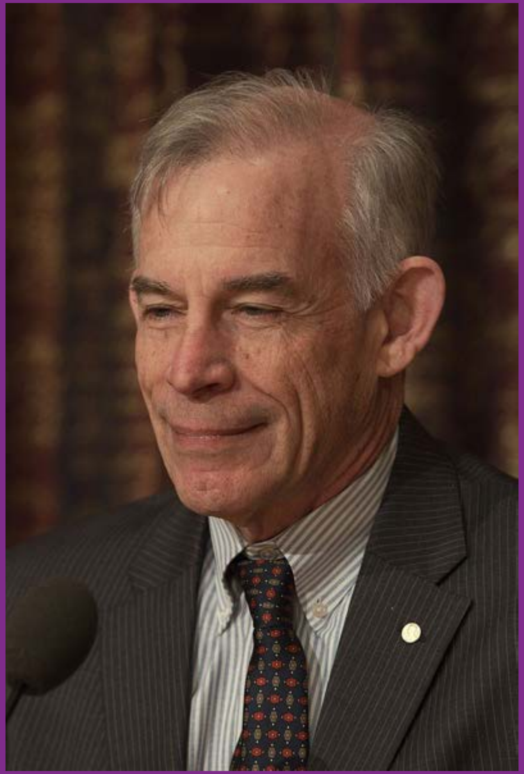
清華大學
Tsinghua University

目录

CONTENTS

1

VAR模型的理论基础



VAR (Auto-Regressions) Model 因Sims（1980）的工作在宏观经济学中成为主流方法， Sims和Sargent获得2011年诺贝尔奖

从ARMA到VAR

※ARMA模型是单变量（Univariate approach）的建模，表现的是同一变量不同时期的相互关系。在现实经济世界中，经济变量相互联系，如价格和成交量，需要考察多个变量不同时期的关系。

※ ARMA可以一般化为VARMA (Vector Autoregressive Moving Average)
缺点：1）随着变量个数N的增加，待估计的参数个数迅速增加，N^2（p+q+1)，极大地增加了估计难度。
2）需要寻找N个高斯白噪音及其合适的移动平均。

※ 解决办法：
1）暴力解决：使用更强大的计算机。不能解决的问题：如何确定q
2）可以证明，一个足够高阶的VAR与VARMA等价。问题简化为估计一个足够高阶的VAR，使得残差值向量为（高斯）白噪音向量。

※ VAR的优缺点：
优点：1）没有内生性问题：互为因果的情况考虑在VAR模型之内
2）估计简便：不但可以使用MLE，每个方程还可以单独使用OLS
3）没有识别问题（辩证地看，也是缺点）：结构模型的识别灾难
缺点：1）系数难以解释（在一个足够高阶的VAR中，有的系数为正，有的为负，有的不显著）
2）缺乏经济理论基础和无法揭示经济变量之间的结构关系（在一个变量较多的VAR中，为什么是这些变量？事物是广泛联系的≠ 事物是乱联系的）一种的解决办法：SVAR
3）估计过程损失较多自由度（尽管优于VARMA，仍然有 N^2p个参数）

VAR的基本设定

※VAR(p):

$$y_t = \upsilon + A_1 y_{t-1} + \ldots + A_p y_{t-p} + u_t$$

其中： $y_t = (y_{1,t}, \ldots, y_{k,t})'$ 为（k*1) 变量向量,必须是平稳时间序列

$\upsilon = (\upsilon_1, \ldots, \upsilon_k)'$ 为（k*1) 常数向量

$A_i, i = 1,2 \ldots k$ 为（k*k)系数矩阵

$u_t = (u_{1,t}, \ldots, u_{k,t})'$ 为（k*1）（高斯）白噪音向量

VAR(p) 是k个方程组成的方程组，每个变量$y_{i,t}$由它自己和其他变量过去p期的取值解释：

$$\begin{aligned} y_{i,t} = v_i &+ \phi_{i,11} y_{1,t-1} + \phi_{i,12} y_{2,t-1} + \cdots + \phi_{i,1k} y_{k,t-1} \\ &+ \phi_{i,21} y_{1,t-2} + \phi_{i,22} y_{2,t-2} + \cdots + \phi_{i,2k} y_{k,t-2} \\ &+ \cdots \\ &+ \phi_{i,k1} y_{1,t-2} + \phi_{i,k2} y_{2,t-2} + \cdots + \phi_{i,kk,} y_{k,t-2} + u_{i,t} \end{aligned}$$

我们可以看到：将VAR（p) 写成方程组的形式比较繁琐，事实上VAR（p) 可以写成VAR(1)的形式：

$$Y_t = \boldsymbol{\upsilon} + \mathbf{A} Y_{t-1} + U_t$$

其中

$$Y_t = \underset{(kp\times 1)}{\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}}, \quad \boldsymbol{\upsilon} = \underset{(kp\times 1)}{\begin{bmatrix} \upsilon \\ 0 \\ \vdots \\ 0 \end{bmatrix}}, \quad U_t = \underset{(kp\times 1)}{\begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}}, \quad \mathbf{A} = \underset{(kp\times kp)}{\begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_k & 0 & \cdots & 0 & 0 \\ 0 & I_k & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_k & 0 \end{bmatrix}}$$

p的确定

※VAR(p):时间序列建模的关键在于截面参数p的确定（Unit root tests, ARMA, Granger causality test, likelihood, ARMA, VAR）

※基本原则：使得VAR模型的残差值为（高斯）白噪音的最小的p的取值

※具体方法：跟我们使用的估计方法有关
方法一：如果我们使用的是MLE，可以用信息准则来判断最优的p：
第一步：确定一个可能的最大的p的取值 $p_{\max}$ （具体情况具体分析如选举周期、生产时长、会计核算等等）
第二步：计算出从1到$p_{\max}$各个模型的信息准则，根据信息准则选择最优滞后阶数p_{optimal}
具体标准：

依据	logL(对数似然函数)	LR(likelihood ratio test)	FPE(final prediction error)	AIC	SIC	HQ
取舍标准	尽可能大	5%水平上显著	最小	最小	最小	最小

方法二：lag exclusion test（LR Test）检验：
第一步：确定一个可能的最大的p的取值$p_{\max}$
H0： $p_{\max}$不显著，H1：H0： $p_{\max}$显著。统计检验量服从χ^2（k^2）分布（证明从略）
如果我们拒绝H0，则使用$p_{\max}$，（注意：此时暗示我们可能需要更高阶的p）,如果我们不能拒绝H0，则转向第二步。
第二步：检验H0： $p_{\max}-1$不显著，H1：H0： $p_{\max}$-1显著
重复第一步，直到不能拒绝原假设

VAR的估计1：MLS

※VAR(p)的估计比ARMA(p,q)简单，既可以使用MLE，也可以使用MLS(Multiple Least Square)

※ MLS：将VAR(p)改写成:Y = BZ + U

$$Y \equiv (y_1, \ldots, y_T)$$
$$B \equiv (\upsilon, A_1, \ldots, A_p)$$
$$Z \equiv (Z_0, \ldots, Z_{T-1})$$
$$U \equiv (u_1, \ldots, u_T)$$

$$Z_t \equiv \begin{bmatrix} 1 \\ y_t \\ \vdots \\ y_{t-p+1} \end{bmatrix}$$

MLS估计量可以直接运用OLS的公式：

$$\mathbf{y} \equiv vec(Y)$$
$$\boldsymbol{\beta} \equiv vec(B)$$
$$\mathbf{u} \equiv vec(U)$$
$$\mathbf{y} = (Z' \otimes I_K)\boldsymbol{\beta} + \mathbf{u}$$
$$\hat{\boldsymbol{\beta}} = ((ZZ')^{-1} \otimes \boldsymbol{\Sigma}_u)(Z \otimes \boldsymbol{\Sigma}_u^{-1})\mathbf{y}$$
$$= ((ZZ')^{-1} Z \otimes I_K)\mathbf{y}$$

其中：⊗为克罗内克积Kronecker product

※优点：随机扰动项只需要是白噪音，无需正态分布

VAR的估计2：MLE

※假定随机扰动项服从正态分布$\boldsymbol{U_t} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Omega})$其中$\boldsymbol{\Omega} = \begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_k^2 \end{pmatrix}$,主对角线上元素可以不一样.

满足这个额外的假定，可以使用MLE

※优点：方面构建一系列的统计检验，如协整中的trace test 和前面的LR lag exclusion test

※缺点：需要对残差值做正态性检验，实际建模中常常难以达到。

最大似然估计的表达式较为冗长，此处从略。

VAR的validation

※VAR的建模步骤同样可以使用Box–Jenkins的四步法：Specification, Estimation, Validation, and Prevision.

Specification: 经济系统的建立（变量的选择），平稳性检验，差分至平稳（若有需要），p的确定（有时需要和第三步残差值检验结合，重新Specification ）

Estimation：选择合适的估计方法（有时需要根据第三步重新估计，如即使一个很高阶的p也无法取得高斯白噪音时，需要退而求其次使用其他估计方法如MLS）。

Validation: 残差值的白噪音检验，正态性检验，异方差性检验，参数的显著性检验（可能存在某种经济结构）

Prevision: 样本内预测/样本外预测,静态一步预测/动态预测，与ARMA预测类似。样本内预测只表达了模型的拟合程度，样本外预测动态与静态相同，样本内预测动态与静态不同。

VAR的prevision

※样本外预测：预测未来h期的各变量取值 $y_{t+h} = (y_{1,t+h}, \ldots, y_{k,t+h})$

预测：

通过迭代的方式向前递推：

$$\begin{aligned} \mathrm{E}_t(y_{t+h}) &= v + A_1\mathrm{E}_t(y_{t+h-1}) + \ldots + A_p\mathrm{E}_t(y_{t+h-p}) + \mathrm{E}_t(u_{t+h}) \\ &= v + A_1\mathrm{E}_t(y_{t+h-1}) + \ldots + A_p\mathrm{E}_t(y_{t+h-p}) \end{aligned}$$

$$\begin{aligned} \mathrm{E}_t(y_{t+1}) &= v + A_1 y_t + \ldots + A_p y_{t-p+1} \\ \mathrm{E}_t(y_{t+2}) &= v + A_1\mathrm{E}_t(y_{t+1}) + A_2 y_t + \ldots + A_p y_{t-p+2} \\ &\vdots \end{aligned}$$

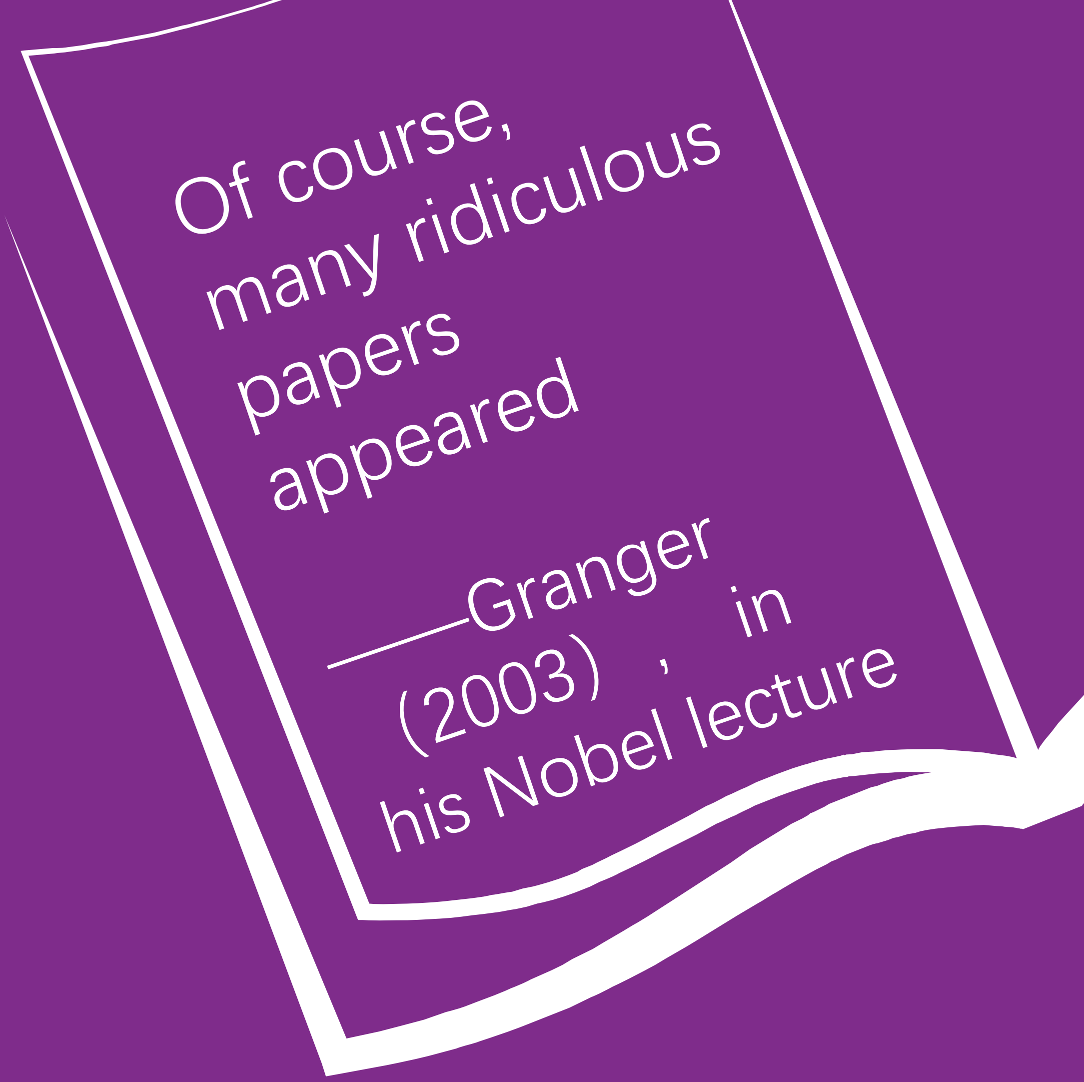
※损失函数(forecast mean squared errors, MSE)：

$$MSE\left[\mathrm{E}_t(y_{t+h})\right] = E\left[(y_{t+h} - \mathrm{E}_t(y_{t+h}))(y_{t+h} - \mathrm{E}_t(y_{t+h}))'\right]$$

基于VAR（p）的预测可以最小化MSE

2

识别、因果检验和结构模型



识别问题

※第一讲和第三讲中提到的内生性问题，在动态结构模型中一般化为识别问题：不同变量不同期的互为因果关系

※问题的起源：Simultaneous equations biases. 早年的（1970s及其以前）宏观计量经济学的实证分析通常基于经济学理论（这段时间主要是凯恩斯主义经济学），提出关于经济结构的实证模型，限定了经济变量之间的特定关系（如IS-LM模型中利率与货币需求等），给出的动态结构模型为Simultaneous equations。然而存在所谓的Simultaneous equations biases：存在多组参数，其概率分布在给定的数据集中是一样的，我们无法区分哪一个是真实的参数。

※凯恩斯主义经济学在1970年代不但遭遇到无法解释的现实经济现象：滞涨现象，学理上也遇到了极大的挑战（计量实证分析和数理理论模型都遭到了全面的批判）。Sims (1980): “incredible identification restrictions” in structural models.凯恩斯主义经济学中大量的实证分析大多存在严重的识别问题。Sims和Sargent的VAR本质上是其理性预期学派的哲学思想。

※一些经济学家（如Blanchard）为恢复凯恩斯主义传统，进行了一系列不懈的努力——计量经济学方面：SVAR（Structural VAR）等，形成了新凯恩斯主义经济学。

※马克思主义经济学家应当向Blanchard等人学习，当现有理论无法对现实做出解释的时候，就应当发展理论，而不是喊口号，空谈信仰！

Granger causality test

※结构模型的识别需要为参数添加某些限制条件，因此可能需要用到“因果检验”。

※ 基本哲学思想：原因早于结果。
在计量经济学的运用：如果X是Y的原因，那么X可以改进Y的预测

※具体方法：假设存在3个信息集：

$$
\begin{aligned}
I_{1,t} &= \{y_{1,t}, y_{1,t-1}, \ldots\} \\
I_{2,t} &= \{y_{2,t}, y_{2,t-1}, \ldots\} \\
I_t &= \{y_{1,t}, y_{1,t-1}, \ldots y_{2,t}, y_{2,t-1}, \ldots\}
\end{aligned}
$$

如果增加变量Y1的信息，对Y2的预测有影响，那么Y1是Y2的Granger因，或者说Y1 Grangerly cause Y2

We say that $y_{1,t}$ Granger-causes $y_{2,t}$ if

$$E\left[y_{2,t} \mid I_{2,t-1}\right] \neq E\left[y_{2,t} \mid I_{t-1}\right]$$

检验设计：在一个VAR(p)模型中：

$$
\begin{aligned}
y_{1,t} &= \Phi_{10} + \Sigma_{i=1}^{p} \Phi_{11}(i) y_{1,t-i} + \Sigma_{i=1}^{p} \Phi_{12}(i) y_{2,t-i} + u_{1,t} \\
y_{2,t} &= \Phi_{20} + \Sigma_{i=1}^{p} \Phi_{21}(i) y_{1,t-i} + \Sigma_{i=1}^{p} \Phi_{22}(i) y_{2,t-i} + u_{2,t}.
\end{aligned}
$$

如果Y1不是Y2的Granger因，显然：$\Phi_{21}(1) = \Phi_{21}(2) = \ldots = \Phi_{21}(p) = 0$
于是检验就简化成了：

$$
\begin{cases}
H_0: & \Phi_{21}(1) = \Phi_{21}(2) = \ldots = \Phi_{21}(p) = 0 \\
H_A: & \Phi_{21}(1) \neq 0 \text{ or } \Phi_{21}(1) \neq 0 \text{ or} \ldots \Phi_{21}(p) \neq 0
\end{cases}
$$

显然约束条件的检验是Fisher检验（整体显著性）或者渐进卡方检验（特定约束条件）

Granger test的意义与局限性

※结构模型的识别需要为参数添加某些限制条件，因此可能需要用到“因果检验”。

※ Granger test只适用于平稳时间序列。

※ Granger test只是统计估计，不是真正意义上的因果关系，不能滥用！所以只是”Grangerly” cause.

※ Granger test结果受其他未考虑在系统中的遗漏变量影响很大，例如，美国的天气状况和中国的大豆油可能没有直接的 Granger causality,但美国的天气如果影响了大豆的收成，进而影响到中国豆油的价格，可能有真实的非直接因果关系。

※ Granger test结果受滞后阶数p的影响极大，很多时候计量结果甚至是被操纵的！

※ Granger test不是万能，只是为我们探索因果关系提供了一种可选的方法，正如Granger 本人说道：” Of course, many ridiculous papers appeared” （2003）， in his Nobel lecture

脉冲响应函数Impulse responses的基本概念

※VAR的系数难以解释，通常没有明确的经济意义。例如，用通常的OLS的观念去看待VAR的系数，无法理解为什么同一变量的不同滞后期的系数，有的为正，有的为负，有的不显著（为0）。（做实证分析的目的是解释经济现象，而不是用经济现象解释模型参数）。

※为什么要脉冲响应函数：在VAR和SVAR中，脉冲响应函数取代了OLS中的参数解释，定量分析某个变量的外部冲击，随着时间的推移，会给系统造成多大的影响。

※一个简单的例子：考虑一个3变量的VAR(1), 简单起见，假定常数项为0.

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix}$$

假定初始值均为0，在t=0时刻，y1变量受到了一个单位的外部冲击

迭代方程，我们发现 $(A_1)^i$ 衡量了i期以后，外部冲击造成的累积影响，因此我们称之为脉冲响应函数。
在VAR(1)的MA表达形式中，它就等于ϕ_i
(Wold's theorem)

$$y_0 = \begin{bmatrix} y_{1,0} \\ y_{2,0} \\ y_{3,0} \end{bmatrix} = \begin{bmatrix} u_{1,0} \\ u_{2,0} \\ u_{3,0} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$y_1 = A_1 y_0$$

$$y_2 = A_1 y_1 = (A_1)^2 y_0$$

$$\cdots$$

脉冲响应函数的一般表达式

※示例的局限性：

1）外部冲击被假定为是孤立的作用于一个变量，多个外部冲击相互影响的情况是怎样呢？

2）VAR(1)的特殊形式，更加一般的VAR(p)是怎样呢？

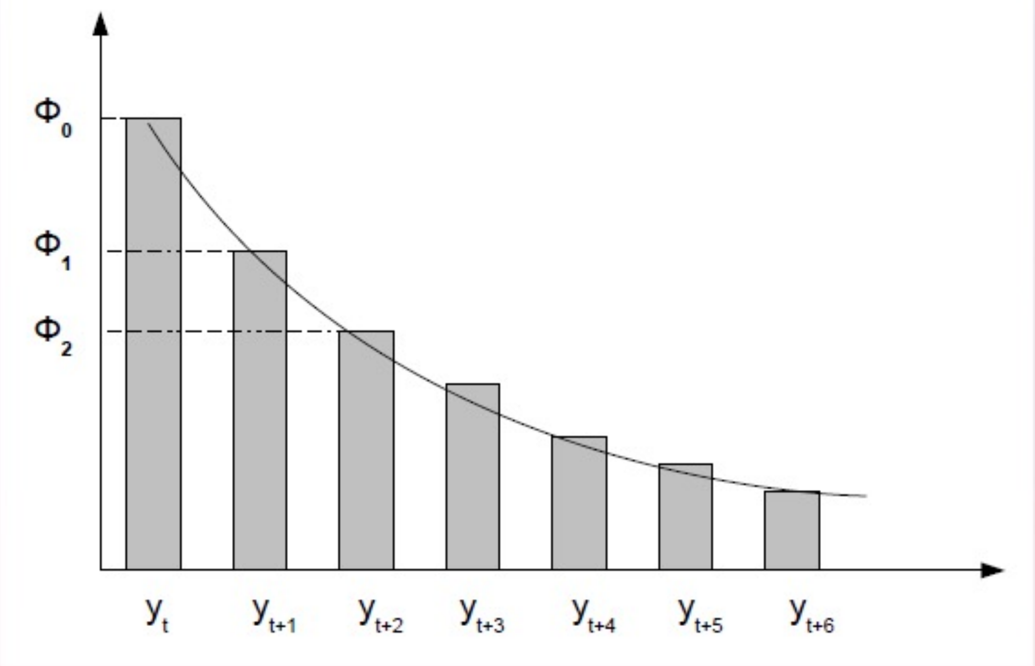
※ Wold ‘s theorem（1954）：任何一个平稳时间序列都可以写为某个白噪音无穷期之和的形式：

$$y_t = \sum_{i=0}^{\infty} \Theta_i w_{t-i}$$

显然脉冲响应函数是 Θ_i

由于VAR（p)总可以写成VAR(1)的形式， VAR（p)的IR可以容易的计算得到。

※ Cholesky分解：多个外部冲击相互依存，需要将其正交分解得到脉冲响应函数（证明从略）



从VAR到SVAR

※为什么要有SVAR?尽管VAR模型可以有效地捕捉到不同变量之间的相互影响，但它无法揭示变量之间的因果机制。有时我们需要了解不同经济变量之前的关系，SVAR提供了这种可能性。

※重要文献： Blanchard and Quah (1989)， Gali（1992），Gerlach and Smets （1995）

※结构向量自回归模型的表达式：

$$B_0 y_t = c_0 + B_1 y_{t-1} + B_2 y_{t-2} + \cdots + B_p y_{t-p} + \epsilon_t,$$

与VAR相比，多了一个结构矩阵B0，不同变量之间存在direct contemporaneous effect

写成reduced form:

$$y_t = B_0^{-1} c_0 + B_0^{-1} B_1 y_{t-1} + B_0^{-1} B_2 y_{t-2} + \cdots + B_0^{-1} B_p y_{t-p} + B_0^{-1} \epsilon_t,$$

我们将SVAR变成了VAR，于是可以应用前面所有关于VAR的已知结论和方法，但我们需要对结构矩阵B0做出一定的限制（假设），模型才是可识别的。（显然，一个三角矩阵是可逆的，这是一种理想的情况）。

$$e_t = B_0^{-1} \varepsilon_t$$

※约束条件：我们需要对B0中的元素做出n(n + 1)/2个约束条件，使得模型是可识别的。

短期约束条件：将B0中某些元素限制为0（结构冲击e_t对内生变量没有当期的影响）

长期约束条件：计算脉冲响应函数，限制某些脉冲响应函数为0，对应于矩阵 $(I + \Phi + \ldots + \Phi^k + \ldots)B$ 中的元素（结构冲击对内生性变量没累积的影响）

VAR的延展

※贝叶斯结构向量自回归模型BSVAR， 见Sims and Zha (2006)

※ Panel vector autoregression, 见Holtz , Newey, and Rosen (1988)

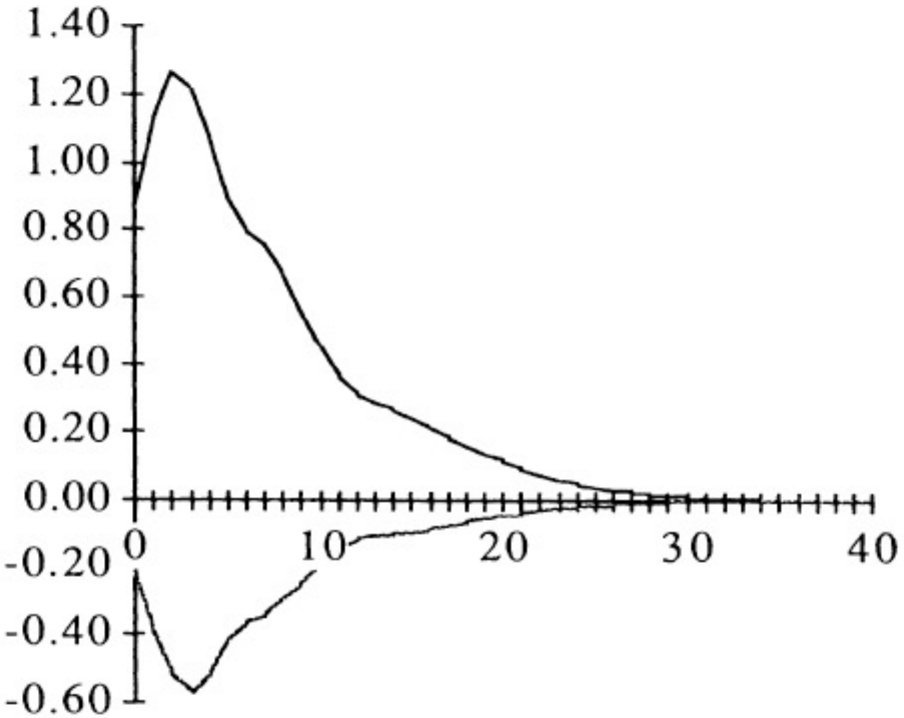


FIGURE 1. RESPONSE TO DEMAND, — = OUTPUT, ~~~ = UNEMPLOYMENT

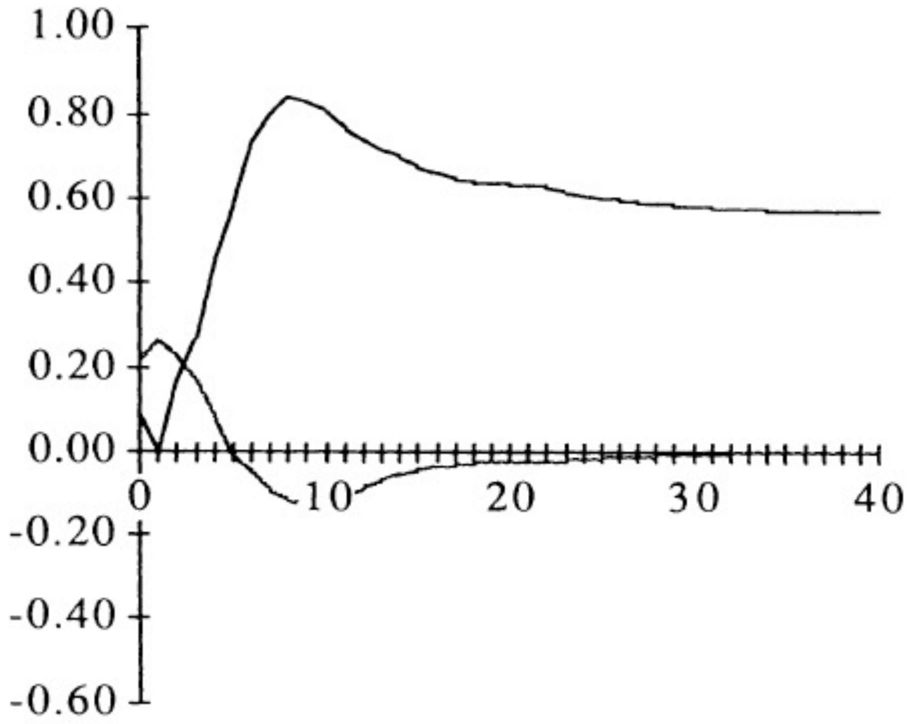


FIGURE 2. RESPONSE TO SUPPLY, — = OUTPUT, ~~~ = UNEMPLOYMENT

3

协整理论

从平稳时间序列建模到非平稳时间序列建模

※平稳时间序列建模的局限性：
1）某些理论模型无法差分，必须使用原序列。
2）平稳时间序列的建模考虑的都是短期关系，无法考察长期关系。

※协整的概念（cointegration）
协整的应用场景：两个序列X和Y是不平稳的；X和Y有共同的随机趋势（共同演变）
例子：上海和全国的物价指数，见LONG and Herrera (2016)
需要总和考虑二者的长期关系和短期关系：
长期模关系：

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

Error correction model（差分后成为平稳时间序列）：

$$\Delta Y_t = a_1 + a_2 \Delta X_t + \pi u_{t-1} + \varepsilon_t$$

协整的定义（简单）：
Yt 和 Xt ~I（d），αXt + βYt~I(d−b), 其中d-b>0,α、β不全为0，那么称Yt 和 Xt ~CI(d,b)
绝大多数金融时间序列都是CI(1，1),也称为 Yt 和 Xt 一阶协整。

※注意：1）时间序列单整阶数必须一致
2）d=b是特殊形式，表明两个非平稳实现序列的非零线性组合是平稳的，最常见的是一阶协整 Yt 和 Xt ~I（1），αXt + βYt~I(0)
3) 协整向量（α，β）可以不止一个

Univariate approach: Engle and Granger (1987)

清華大學 Tsinghua University

※Engle and Granger (1987)针对两个变量的协整检验提出了以下4步法：

第一步：检验Yt 和 Xt 的平稳性，　若Yt 和 Xt ~I(0), 使用标准的平稳时间序列建模

若Yt 和 Xt 单整阶数不同，它们不是协整的

若Yt 和 Xt 单整阶数相同，转向第二步

第二步：估计方程的OLS估计量：　$Y_t = \beta_1 + \beta_2 X_t + e_t$

计算残差值序列$\hat{e}_t$.

Stock (1987) 证明，如果残差值是平稳的，那么OLS估计量是超收敛的（Super-convergent）

如果残差值不平稳，则是伪回归

第三步：检验残差值序列的平稳性

由于时间序列是不平稳的，**残差值不是随机扰动项的无偏差估计量**，而是有偏估计量，常用的ADF检验临界值并不能直接用于残差值的平稳性检验（包括**Eviews和Stata等软件直接给出的临界值都不可以直接用**）需要使用Engle and Granger (1987)给出的临界值表：

	1%	5%	10%
No lags	-4,07	-3,37	-3,30
Lags	-3,73	-3,17	-2,91

许多统计学教材都没有搞明白这一点，《经济研究》上有大量论文也都在这个问题上犯错误，许多论文都是错的！

其他检验：Phillips and Ouliaris (1990) 给出了Phillips-Perron test 的方法

第四步：使用OLS估计EMC：$\Delta Y_t = a_0 + b_1 \Delta X_t + \pi \hat{e}_{t-1} + v_t$ 其中：$\hat{e}_{t-1} = Y_{t-1} - \hat{\beta}_1 - \hat{\beta}_2 X_{t-1}$

临界值表

$$\hat{C}(p,T) = \hat{\beta}_\infty + \beta_1 T^{-1} + \beta_2 T^{-2}$$

p = sig. lev.

TABLE I. Response surface estimates of critical values

N	Variant	Size (%)	Obs.	β_∞	(SE)	β_1	β_2
1	No constant	1	600	−2.5658	(.0023)	−1.960	−10.04
		5	600	−1.9393	(.0008)	−0.398	0.0
		10	560	−1.6156	(.0007)	−0.181	0.0
1	No trend	1	600	−3.4335	(.0024)	−5.999	−29.25
		5	600	−2.8621	(.0011)	−2.738	−8.36
		10	600	−2.5671	(.0009)	−1.438	−4.48
1	With trend	1	600	−3.9638	(.0019)	−8.353	−47.44
		5	600	−3.4126	(.0012)	−4.039	−17.83
		10	600	−3.1279	(.0009)	−2.418	−7.58
2	No trend	1	600	−3.9001	(.0022)	−10.534	−30.03
		5	600	−3.3377	(.0012)	−5.967	−8.98
		10	600	−3.0462	(.0009)	−4.069	−5.73
2	With trend	1	600	−4.3266	(.0022)	−15.531	−34.03
		5	560	−3.7809	(.0013)	−9.421	−15.06
		10	600	−3.4959	(.0009)	−7.203	−4.01
3	No trend	1	560	−4.2981	(.0023)	−13.790	−46.37
		5	560	−3.7429	(.0012)	−8.352	−13.41
		10	600	−3.4518	(.0010)	−6.241	−2.79
3	With trend	1	600	−4.6676	(.0022)	−18.492	−49.35
		5	600	−4.1193	(.0011)	−12.024	−13.13
		10	600	−3.8344	(.0009)	−9.188	−4.85
4	No trend	1	560	−4.6493	(.0023)	−17.188	−59.20
		5	560	−4.1000	(.0012)	−10.745	−21.57
		10	600	−3.8110	(.0009)	−8.317	−5.19
4	With trend	1	600	−4.9695	(.0021)	−22.504	−50.22
		5	560	−4.4294	(.0012)	−14.501	−19.54
		10	560	−4.1474	(.0010)	−11.165	−9.88
5	No trend	1	520	−4.9587	(.0026)	−22.140	−37.29
		5	560	−4.4185	(.0013)	−13.641	−21.16
		10	600	−4.1327	(.0009)	−10.638	−5.48
5	With trend	1	600	−5.2497	(.0024)	−26.606	−49.56
		5	600	−4.7154	(.0013)	−17.432	−16.50
		10	600	−4.4345	(.0010)	−13.654	−5.77
6	No trend	1	480	−5.2400	(.0029)	−26.278	−41.65
		5	480	−4.7048	(.0018)	−17.120	−11.17
		10	480	−4.4242	(.0010)	−13.347	0.0
6	With trend	1	480	−5.5127	(.0033)	−30.735	−52.50
		5	480	−4.9767	(.0017)	−20.883	−9.05
		[illegible]	[illegible]	[illegible]	(.0011)	−16.445	0.0

Univariate approach的注释

※EMC中，我们预期调整参数$\pi < 0$.因为，若当$\widehat{e_{t-1}} > 0$，意味着Y_{t-1}在其均衡水平上($Y_{t-1} > \hat{\beta}_1 + \hat{\beta}_2 X_{t-1}$)需要通过$\Delta Y_t < 0$来修正.

※ Phillips (1986)证明，长期模型的OLS估计量不服从渐进正态分布，不适合用来检验显著性。需要使用Fully Modi ed OLS (Phillips and Hansen, 1990) 或者Canonical Cointegrating Regression (CCR) (Park, 1992)

※在样本量足够大的时候，$Y_t = \beta_1 + \beta_2 X_t + e_{1,t}$ 和 $X_t = \beta_1 + \beta_2 Y_t + e_{2,t}$ 的协整检验结论是一样的，但在有限样本中，结论并不一致，多变量时更加复杂。

※ Engle-Granger Univariate approach只能考虑一个协整向量，然而在多个变量的时候，可能存在多个协整向量。

※ Engle-Granger Univariate approach 分为四步，每一步有错误都会导致下一步出错
其他办法：Johansen multivariate approach.

Johansen multivariate approach的基本方法

※类似于ARMA扩展为VAR，EMC扩展为VEMC，可以很容易证明（从略）一个VAR(p)等价于VEMC(p-1)

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t$$

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p-1} + u_t$$

$$\Pi = -(I_k - A_1 - \dots - A_p) \qquad \Gamma_i = -(A_{i+1} + \dots + A_p)$$

※协整检验变为检验矩阵Π（k*k）的秩r(0<r<=k-1)：
若Π满秩，即r=k, Yt是平稳的
若rang(Π)=0, 不存在协整关系

※估计方法：MLE

※注释：
Π表达了长期关系，因此Π的秩是多少，就给出了多少个协整关系
Γ_i是短期关系
协整检验直接由trace test 或eigen value test 给出

Johansen multivariate approach的步骤1

类似于Univariate approach, Johansen multivariate approach分为6个步骤：

※第一步： 检验所有内生变量的平稳性

※第二步：选择合适的p估计VAR（p）

※第三步：选择长期模型部分中合适的趋势项类型：

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p-1} + \alpha \begin{pmatrix} \beta \\ \mu_1 \\ \delta_1 \end{pmatrix} (y_{t-1}\ 1\ t)$$
$$+\mu_2 + \delta_2 t + u_t,$$

五种类型：

a) 协整方差（CE）和VAR均中没有常数项和趋势项 $\mu_1 = \delta_1 = \mu_2 = \delta_2 = 0$

不太可能,因为一般来说不同变量之间测量单位不同，CE都会有常数，

b) 协整方差（CE）中有常数项，VAR中没有常数项和趋势项 $\delta_1 = \mu_2 = \delta_2 = 0$

$\Delta Y_t = 0$

c) CE和VAR中有常数项，但没有趋势项 $\delta_1 = \delta_2 = 0$

$\Delta Y_t \neq 0$但时间序列没有趋势项

d) CE中有常数项和趋势项，VAR中没有趋势项 $\delta_2 = 0.$

表明外部因素影响变量的增长（如技术进步）

e) CE和VAR中均有趋势项和常数项

罕见，从经济上很难解释，因为意味着递增或递减的增速

Johansen multivariate approach的步骤2

第三步的选择标准：
基本原则：根据经济意义选择
参考性原则：Pantula principle (Hansen and Juselius (1995)) :(Hjelm and Johansson (2005))
从最严格的H0(a类情况0个协整关系)到最宽松的H0（e类情况，k-1个协整关系）
经验性原则：通常在b,c,d中间选

※第四步：确定协整关系的个数
最大特征值检验
Trace test (更有效，推荐使用)

※第五步（选）：检验弱外生性

※第六步（选）：检验协整向量的线性约束

4

参考文献


站在巨人的肩膀上
"If I have seen further it is by standing on the shoulders of Giants. "
by Isaac Newton in 1675

参考文献&数学基础知识&样本数据

见网络学堂附件

下节课见

马克思主义学院
龙治铭

清華大學
Tsinghua University