第十二讲 协整和 ECM 模型

概要：本讲继续分析两个变量的时间序列，一个输入变量$\{x_t\}$，一个输出变量$\{y_t\}$，且这两个变量是非平稳的。非平稳变量构建动态回归模型，会产生伪回归，其参数有效性检验结果不可信，在这种情况下，引入协整概念，即非平稳变量的回归残差序列为平稳的时间序列，说明两个非平稳变量间存在稳定的长期关系。在构建协整模型的基础上，需要构建 ECM 模型（误差修正模型），解释短期波动产生的原因。此外，宏观经济变量之间都是相互影响的，在实际建模时，需要采用 Granger 因果检验确定输入变量（自变量）和输出变量（因变量）。

无论是一元时间序列分析还是多元时间序列分析，构建模型的前提条件都是平稳。在第九讲中，讲解了结构最简单的平稳双变量时间序列模型，转换函数模型。在本讲中，继续讲解双变量模型，只是了两个变量都是非平稳的。对于包含非平稳变量的处理，在双变量模型中，可以通过差分去除趋势和周期，从而生成平稳序列，然后采用转换函数模型的方法。如果将非平稳变量直接放入动态回归模型中进行建模，会产生“伪回归”，此外，非平稳变量的线性组合很可能是平稳的，这些变量被称为具有协整关系 Cointegration；协整关系反映了变量之间长期关系，即构建长期模型，在此基础上，可以同时构建变量之间的短期关系模型，即 ECM 模型（Error Correction Model 误差修正模型）。许多宏观经济分析中，都包含这种协整分析（长期）和 ECM 分析[1]（短期）。

一、伪回归 Spurious Regression

考虑双变量回归方程：

$$y_t = a_0 + a_1 x_t + \varepsilon_t \qquad 10.1$$

$\{y_t\}$和$\{x_t\}$是两个序列相关的时间序列（非白噪声），ε_t代表的是误差项，$\{\varepsilon_t\}$也构成时间序列，存在着序列相关（非白噪声）。要采用式 10.1 的古典回归模型，$\{y_t\}$和$\{x_t\}$必须是平稳时序，$\{\varepsilon_t\}$的均值为 0，即$E(\varepsilon_t) = 0$，方差有界，即$Var(\varepsilon_t) < \infty$。如果进入回归模型的

[1] 只有一个方程的称为 ECM 模型，Error Correction Model，误差修正模型；包含两个及以上方程的称为 VEC 模型或 VECM 模型，Vector Error Correction Model，向量误差下修正模型

变量是非平稳的，那么就会造成“伪回归”，即在参数估计后的模型检验中，得到的R^2值（Goodness of fit，拟合优度）很高，t值显著，但回归结果无任何经济意义，或者说，回归结果看上去很好，但 OLS 估计量不一致，且常规的t检验和F检验都失效。Granger & Newbold（1974）采用以下方法考察了“伪回归”的产生及后果：

第一步，采用数据模型的方法，分别形成两个独立的随机游走 Random Walk 序列$\{y_t\}$和$\{z_t\}$，这两个序列是典型的非平稳序列。即：

$$y_t = y_{t-1} + e_{yt} \qquad 10.2$$

$$x_t = x_{t-1} + e_{xt} \qquad 10.3$$

其中，e_{yt}和e_{xt}是白噪声序列，且相互独立。

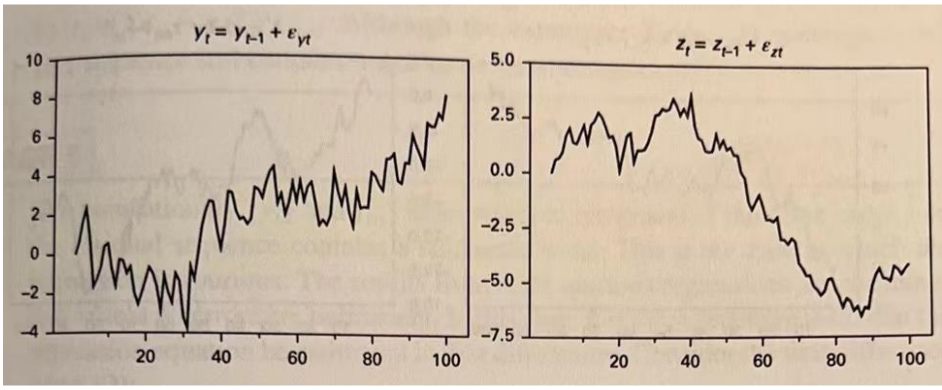
第二步，采用式 10.2 和式 10.3 形成的数据，以式 10.1 的方式进行回归。由于$\{y_t\}$和$\{x_t\}$这两个序列的观测值是分别形成的，它们之间是相互独立的。虽然运行软件可以得到式 10.1 的回归方程，但此方程毫无意义的，任何描绘这两个时序变量关系的模型都是“假的，伪的”spurious。

第三步，重复第一步和第二步 1000 次，形成 1000 个形如式 10.1 的回归方程，观察这 1000 个方程参数a_1的t值分布表（见表 10.1），发现：t值（绝对值）在 2 以上的回归方程为 675 个，在 5%的显著性水平下（t 值为 1.96），有 700 多个方程会拒绝H_0：$a_1 = 0$的假设，而按正常情况，这个数字应该在 50 个方程，说明$\{y_t\}$和$\{x_t\}$为非平稳时，构建的回归模型在假设检验时失败，样本方程无法有效说明总体方程。

表 10.1：$y_t = a_0 + a_1 x_t + \varepsilon_t$模型的 t 值（绝对值）频度分布表

（y_t和x_t均为一阶非平稳过程）

t	0–2	2–4	4–6	6–8	8–10	10–12	12–14	14–16	16–18	>18
Frequency	325	281	178	98	67	27	15	3	4	2



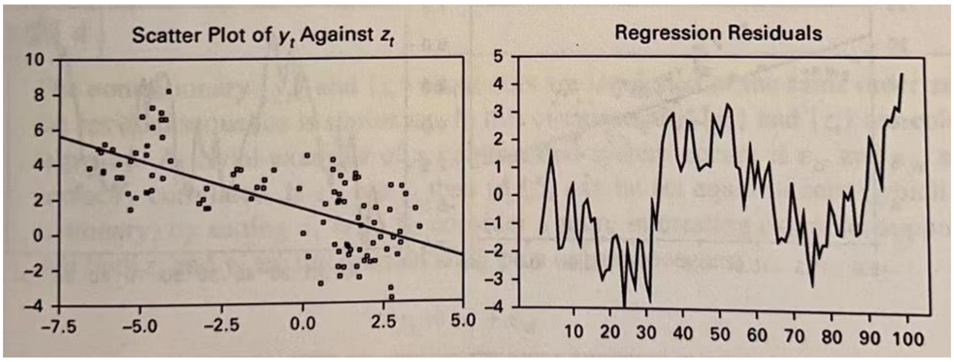


图 10.1，“伪回归”过程（图中的z_t就是x_t）

图 10.1 的四张图给出了“伪回归”的形成过程。上图分别是随机游走$\{y_t\}$和$\{x_t\}$的时序图，其中，$\{y_t\}$是向右上方倾斜，$\{x_t\}$是向右下方倾斜，经计算，二者之间的相关系数是-0.69，中度负相关；下图左边是$\{y_t\}$和$\{x_t\}$的散点图，以及经过 OLS 回归得到的直线方程$y_t = 1.41 - 0.565x_t$，下图右边是回归的残差图，可以看出是一个典型的非平稳序列。实际上，$\{y_t\}$和$\{x_t\}$这两个非平稳时序是通过模型方法分别形成的，之间毫无联系。

二、单整和协整

2.1 研究背景

对于货币市场的研究促进了单整和协整概念及相关模型的发展。

首先考虑一个简单的货币需求模型。理论表明，个人希望持有的是真实货币（real money），而在现实中，大家所拥有的是名义货币（nominal money），后者的需求与价格水平成正比，也就是说，物价水平越高，所需要的名义货币就越多；此外，随着收入水平和相关交易数量的增加，个人将希望持有更多的名义货币；进一步，由于利率是持有货币的最大机会成本，名义货币需求应该与利率负相关。上述的理论可以写成一个经济计量模型：

$$m_t = \beta_0 + \beta_1 p_t + \beta_2 y_t + \beta_3 r_t + \varepsilon_t \qquad 10.4$$

其中：m_t为货币需求，p_t为价格水平，y_t为真实收入水平，r_t为利率水平，ε_t为干扰项，$\beta_0, \beta_1, \beta_2, \beta_3$为待估参数。式中，除了利率$r_t$外，其余变量均取对数形式（log）。

假定货币市场处于均衡状态，此时，名义货币量等于货币供给量。此时，可以收集货币供给量、价格水平、真实收入水平（可以用真实 GDP 来衡量）和短期利率的时间序列数据，构建式 10.4 模型，该模型要求$\beta_1 = 1$， $\beta_2 > 0$，且$\beta_3 < 0$。除了上述三个变量影响货币需求外，模型中未解释部分的性质（即，$\{e_t\}$序列）也很重要，如果理论的均衡假设是成立的，

那么货币需求的任何偏差必然是暂时的，也就是说，在时间推移的过程中，这种偏差性的波动不会消失，但如果均衡存在，这个偏差将是收敛的，即，理论的均衡假设要求$\{e_t\}$序列是平稳的。通常，真实 GDP、货币供应量（名义货币需求量）、价格水平和利率都是非平稳的时间序列，每个变量都可以蜿蜒变化而回归到长期水平。而式 10.4 模型所表达的意思是，这些非平稳变量的线性组合是平稳的。我们可以通过重写式 10.4 为：

$$\varepsilon_t = m_t - \beta_0 - \beta_1 p_t - \beta_2 y_t - \beta_3 r_t \qquad 10.5$$

由于$\{\varepsilon_t\}$序列是平稳的，那么，式 10.5 的右边，四个变量的线性组合也必须是平稳的。也就是说，均衡理论要求随着时间的变动，四个非平稳变量$\{m_t\}$、$\{p_t\}$、$\{y_t\}$和$\{r_t\}$之间的变动是相互关联的。上述例子揭示了宏观计量模型的关键见解：均衡理论要求非平稳变量的组合是平稳的。许多宏观经济模型也是如此。

2.2 单整的概念

为了构建多变量非平稳时序模型，首先给出了单整的概念。单整概念就是非平稳概念的另一套说法。

在单位根检验的过程中，如果检验结果显著拒绝序列非平稳的原假设，即说明$\{x_t\}$**显著平稳，不存在单位根，这时称序列**$\{x_t\}$**为零阶单整**（integration）序列，简记为$x_t \sim I(0)$。假如原假设不能被显著拒绝，说明序列$\{x_t\}$为非平稳序列，存在单位根。这时可以考虑对该序列进行适当阶数的差分，以消除单位根，实现平稳。**如果原序列一阶差分后平稳，说明原序列存在一个单位根，这时称原序列为 1 阶单整序列，简记为**$x_t \sim I(1)$。如果原序列至少需要进行 d 阶差分后才能实现平稳，说明原序列存在 d 个单位根，这时候称原序列为 d 阶单整序列，简记为$x_t \sim I(d)$。

单整衡量的是单个序列的平稳性，它具有如下重要性质：

（1）若$x_t \sim I(0)$，对于任意非零实数a, b，有$a + bx_t \sim I(0)$；

（2）若$x_t \sim I(d)$，对于任意非零实数a, b，有$a + bx_t \sim I(d)$；

（3）若$x_t \sim I(0), y_t \sim I(0)$，对于任意非零实数$a, b$，有$ax_t + by_t \sim I(0)$；

（4）若$x_t \sim I(d), y_t \sim I(c)$，对于任意非零实数$a, b$，有$ax_t + by_t \sim I(k)$，$k \le \max[d, c]$

2.3 协整的概念

从 2.1 的叙述中可知，有些序列自身的变化虽然是非平稳的，但是序列与序列之间却具有非常密切的长期均衡关系。为了有效衡量序列之间是否具有长期均衡关系，Engle 和 Granger 于 1987 年提出协整 Cointegration 这个概念。具体的定义如下：考虑有一系列的经济

变量，经过长期的相互作用达到均衡，即：

$$\beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \cdots + \beta_n x_{nt} = 0 \qquad 10.6$$

令：$\mathbf{B} = (\beta_1, \beta_2, \dots, \beta_n)$，$\mathbf{X}_t = (x_{1t}, x_{2t}, \dots, x_{nt})'$ ，则$\mathbf{BX}_t = 0$ ，说明系统处于长期均衡状态。设定$\boldsymbol{\varepsilon_t}$度量偏离长期均衡的误差 equilibrium error，则模型可以写成：

$$\boldsymbol{\varepsilon_t} = \mathbf{BX}_t \qquad 10.7$$

如果长期均衡存在，则均衡误差$\{\boldsymbol{\varepsilon_t}\}$为平稳序列，即零阶单整 *I(0)*，也就是说，虽然有所偏离，但总是在 0 附近徘徊，均衡是存在的。如果$\boldsymbol{\varepsilon_t} \sim I(1)$，则说明偏离 0 越来越远，不存在均衡（注意：此处“均衡”Equilibrium 的含义与经济学中的含义是有区别的）。

Engle & Granger（1987）对“协整 Cointegration”的定义如下：

列向量$\mathbf{X}_t = (x_{1t}, x_{2t}, \dots, x_{nt})'$ 若满足以下两个条件，则视为存在协整关系，1）向量中的每个变量都是d阶单整，即$I(d)$；2）存在协整向量 Cointegrating Vector $\mathbf{B} = (\beta_1, \beta_2, \dots, \beta_n)$，使得线性组合$\mathbf{BX}_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \cdots + \beta_n x_{nt}$为$(d-b)$阶单整，其中$b > 0$。此时，协整向量$\mathbf{X}_t$记作$\mathbf{X}_t \sim CI(d, b)$。依此定义，式 4 可以理解为，如果四个变量，货币需求m_t、价格水平p_t、真实收入水平y_t、利率水平r_t均为一阶单整序列 *I(1)*，并且其线性组合$m_t - \beta_0 - \beta_1 p_t - \beta_2 y_t - \beta_3 r_t = \varepsilon_t$是平稳的，则向量$\mathbf{X}_t = (m_t, 1, p_t, y_t, r_t)'$ 存在协整关系，记作$CI(1,1)$，协整向量$\mathbf{B} = (1, -\beta_0, -\beta_1，-\beta_2, -\beta_3)$。$\varepsilon_t$表示偏离货币市场长期均衡的程度，$\{\varepsilon_t\}$为平稳过程，这种偏离都是暂时的。

对于协整 Cointegration 的理解，要注意以下两点：

1）协整关系是线性关系，即非平稳变量的线性组合，非线性关系不在讨论范围之内。此外，协整向量 Cointegrating Vector $\mathbf{B} = (\beta_1, \beta_2, \dots, \beta_n)$并非唯一的，对于任意一个非零数$\lambda$，$(\lambda\beta_1, \lambda\beta_2, \dots, \lambda\beta_n)$都是协整向量；

2）依据 Engle & Granger 最初的定义，只有同阶单整的变量才有协整关系，但这并不意味着任何同阶单整的变量一定存在协整关系，或者说，不存在协整关系的同阶单整变量之间不存在着长期均衡关系，这些变量相互之间蜿蜒徘徊的距离会越来越远。如果两个变量为不同阶单整，则不存在协整关系，例如：$x_{1t} \sim I(d_1)$，$x_{2t} \sim I(d_2)$，并且$d_2 > d_1$，可以证明，x_{1t}和x_{2t}的线性组合是d_2阶单整，$I(d_2)$，即非平稳。

但是，两个以上非同阶单整的变量也有可能存在均衡关系，或协整。例如：$x_{1t} \sim I(2)$，$x_{2t} \sim I(2)$， $x_{3t} \sim I(1)$，x_{1t}或x_{2t}与x_{3t}之间肯定不存在协整关系，但是，如果x_{1t}和x_{2t}的线性组合的关系是$CI(2,1)$，也就是说$\beta_1 x_{1t} + \beta_2 x_{2t}$的结果是 1 阶单整$I(1)$，那么这个线性组合与

x_{3t}之间就有可能构成协整关系。Lee & Granger（1990）称这种为多元协整 Multi-cointegration。

2.3 协整模型

本节以输入变量（自变量）$\{x_t\}$和输出变量$\{y_t\}$（因变量）来说明协整模型的构建过程。

对$\{x_t\}$和$\{y_t\}$进行平稳性检验，如果都是非平稳序列，且为同阶单整序列，才能构建协整模型。

两个非平稳序列之间是否能建立动态回归模型，关键在于它们之间是否存在协整关系。所以多元非平稳序列建模必须先进行协整检验，也称 Engle-Granger 检验，简称 EG 检验。

EG 检验的假设条件为：

H_0: 多元序列之间不存在协整关系　　H_1: 多元序列之间存在协整关系

由于协整关系主要是通过考察残差序列的平稳性，所以上述假设条件等价于：

H_0: 回归残差序列$\{\varepsilon_t\}$非平稳　　H_1: 回归残差序列$\{\varepsilon_t\}$平稳

具体操作步骤如下：

步骤一：建立输出序列和输入序列之间的回归模型：

$$y_t = \widehat{\beta_o} + \widehat{\beta_1} x_t + \varepsilon_t \qquad 10.8$$

步骤二：对回归残差序列$\{\varepsilon_t\}$进行平稳性检验。主要采用单位根检验的方法来考察回归残差的平稳性，所以假设条件等价于：

$$H_0: \ \varepsilon_t \sim I(k) \quad k \geq 1 \qquad\qquad H_0: \ \varepsilon_t \sim I(0)$$

EG 检验原理和计算公式与 ADF 检验原理和计算公式相同，但临界值略有差异。EG 检验的临界值不仅与漂移项、趋势项等因素有关，还与回归模型中非平稳变量的个数有关。当非平稳序列的个数为 1 时，对应的就是 ADF 检验，当非平稳序列的个数大于等于 2 时，对应的就是 EG 检验。

如果回归残差序列$\{\varepsilon_t\}$通过平稳性检验，即$\varepsilon_t \sim I(0)$，就说明输入序列$\{x_t\}$和输出序列$\{y_t\}$之间存在着协整关系，说明两个时序之间存在长期均衡关系，而且这个关系可以用 EG 检验第一步建立的回归模型表达：

$$y_t = \widehat{\beta_o} + \widehat{\beta_1} x_t + \varepsilon_t$$

回归残差序列就是：

$$\varepsilon_t = y_t - \left(\widehat{\beta_o} + \widehat{\beta_1} x_t\right) \qquad 10.9$$

式 10.9 包含输出序列不能由输入序列解释的随机波动。这个随机波动里可能还蕴涵着历史信息之间的相关性，所以进一步考察$\{\varepsilon_t\}$的自相关信息和偏自相关信息，构建 ARMA 模型：

$$\varepsilon_t = \frac{\Theta(B)}{\Phi(B)}\alpha_t \qquad 10.10$$

式 10.10 中，$\Theta(B)$为移动平均系数多项式，$\Phi(B)$为自回归系数多项式，α_t为正态分布的白噪声序列。

完成以上分析，就可以得到输入序列和输出序列的协整拟合模型：

$$y_t = \widehat{\beta_o} + \widehat{\beta_1}x_t + \frac{\Theta(B)}{\Phi(B)}\alpha_t \qquad 10.11$$

三、误差修正模型 ECM

误差修正模型 Error Correction Model，简称 ECM 模型，它常常作为协整模型的补充模型出现。协整模型度量的是序列之间的长期关系，而 ECM 模型则解释了序列的短期波动关系。为了有效理解 ECM，先从**自回归分布滞后模型** Auto Regression Distributed Lag Model，ARDL Model 说起。

ADSL 模型可以视作是转换函数模型的扩展。以双变量的转换函数模型为例：

$$y_t = v_0x_t + v_1x_{t-1} + \cdots + v_nx_{t-n} + \varepsilon_t \qquad 10.12$$

式 10.22 代表双变量的转换函数模型，x_t对y_t的影响要没有延迟，有效作用期长为 *n* 期。现实情况还可能是，输入变量y_t除了受输入变量x_t的影响外，还受到自身的影响，也就是$y_{t-1}, y_{t-2}, \ldots$,的影响，如果输出变量历史值的影响作用时间是 *m* 期，则式 10.12 就扩展成：

$$y_t = \omega_1y_{t-1}+\omega_2y_{t-2} + \cdots + \omega_my_{t-m} + v_0x_t + v_1x_{t-1} + \cdots + v_nx_{t-n} + \varepsilon_t \qquad 10.13$$

式 10.13 就是**自回归分布滞后模型**，ARDL(*m,n*)。ARDL(1,1)可以写成：

$$y_t = \omega_1y_{t-1} + v_0x_t + v_1x_{t-1} + \varepsilon_t \qquad 10.14$$

对式 10.14 做一阶差分，得到：

$$\nabla y_t = (\omega_1 - 1)y_{t-1} + (v_0 + v_1)x_{t-1} + v_0\nabla x_t + \varepsilon_t \qquad 10.15$$

移项整理得到：

$$\nabla y_t = v_0\nabla x_t - (1 - \omega_1)[y_{t-1} - \frac{v_0 + v_1}{1 - \omega_1}x_{t-1}] + \varepsilon_t \qquad 10.16$$

式 10.16 说明输出变量的当下变动∇y_t，是受三个方面的短期波动影响：1）输入变量的当下变动∇x_t；2）上一期偏离“长期均衡”的数量$y_{t-1} - \frac{v_0+v_1}{1-\omega_1}x_{t-1}$，即上一期的误差；3）当期的随机波动$\varepsilon_t$，按教材上的写法，就是：

$$\nabla y_t = \beta_0\nabla x_t + \beta_1 ECM_{t-1} + \varepsilon_t \qquad 10.17$$

式 10.17 和式 10.16 比较，$\beta_0 = v_0$，$\beta_1 = -(1 - \omega_1)$

如果长期均衡模型$y_{t-1}-\frac{v_0+v_1}{1-\omega_1}x_{t-1}$的参数未知，说明$ECM_{t-1}$的值是未知的，式 10.17 的参数估计要采用非线性最小二乘法 Nonlinear Least Squares。如果将式 10.16 改写成：

$$\nabla y_t = v_0 \nabla x_t - (1-\omega_1)[y_{t-1}-x_{t-1}] + (1-\omega_1+v_0+v_1)x_{t-1} + \varepsilon_t \qquad 10.18$$

式 10.18 就可以直接采用最小二乘估计 Ordinary Least Squares。

式 10.17 中的β_1，式 10.16 和式 10.18 中的$-(1-\omega_1)$为**误差修正系数**，表示误差修正项对当期波动的修正力度。根据误差修正系数的推导原理，可以确定$\beta_1 < 0$，$-(1-\omega_1) < 0$，即误差修正机制是一个负反馈机制。对式 10.17 而言，当$ECM_{t-1} > 0$时，说明$y_{t-1} > \widehat{y_{t-1}}$，即真实值比估计值大，这种信息反馈回来，上一期出超会导致下一期适当压缩，即$\nabla y_t < 0$。

四、Granger 因果检验

对于多元时间序列而言，如果能找到对输出序列有显著影响的输入序列，并且能够验证它们之间具有协整关系，就说明输出序列$\{y_t\}$的一部分波动能被输入序列$\{x_t\}$的线性组合所解释。这对于预测$\{y_t\}$的波动，或者通过控制输入序列$\{x_t\}$的取值，间接控制$\{y_t\}$的发展都是非常有用的。但前提是输入序列和输出序列之间具有真正的因果关系，而且一定是$\{x_t\}$为因，$\{y_t\}$为果。这种因果关系的确定，在某些情况下是清晰的，比如：收入影响支出，收入一定是因，支出一定是果。但许多情况，这种情况很难直观判断，特别对于经济变量，大部分情况下是互为因果的，比如：GDP 和国际原油价格，它们之间的因果关系并步一目了然。因此，**在协整建模时，首先需要检验变量之间的因果关系**。在时序分析中，最为常用的是 Granger 因果检验。

4.1 因果关系定义

因果关系，一定是原因导致结果。所以从时间上讲，应该是原因发生在前，结果产生在后。就影响效果而言，X 事件发生在前，而且对 Y 事件的发生结果产生影响，X 事件才能称为 Y 事件的因。如果 X 事件发生在前，但它发生与否对 Y 事件的结果没有影响，X 事件也不是 Y 事件的因。基于对这种因果关系的理解，Granger 给出了序列因果关系的定义，即：

假设$\{x_t\}$和$\{y_t\}$是宽平稳序列，记

（1）I_t为 t 时刻所有有用信息的集合，$I_t = \{x_t, x_{t-1}, x_{t-2}, \ldots, y_t, y_{t-1}, y_{t-2}, \ldots\}$

（2）X_t为 t 时刻所有x_t信息的集合，$X_t = \{x_t, x_{t-1}, x_{t-2}, \ldots\}$

（3）$\sigma^2(\,)$为方差函数

当且仅当y_t的最优线性预测函数使得$\sigma^2(y_{t+1}|I_t) < \sigma^2(y_{t+1}|I_t - X_t)$成立，则$\{x_t\}$是$\{y_t\}$的 Granger 原因。

式中，$\sigma^2(y_{t+1}|I_t)$是使用所有可获得的历史信息（其中也包含$\{x_t\}$的信息）所得到的$\{y_t\}$的 1 期预测值的方差；$\sigma^2(y_{t+1}|I_t - X_t)$是从所有信息中刻意扣除$\{x_t\}$的历史信息得到的$\{y_t\}$的 1 期预测值的方差。如果$\sigma^2(y_{t+1}|I_t) < \sigma^2(y_{t+1}|I_t - X_t)$成立，说明$\{x_t\}$历史信息的加入能提高$\{y_t\}$的预测精度，由此推断，$\{x_t\}$是因，$\{y_t\}$是果，记作：$x_t \to y_t$。

根据 Granger 因果关系的定义，在两个序列之间存在 4 种不同的因果关系（在此不考虑x_{t+1}对y_{t+1}当期影响）：

（1）x_t和y_t相互独立，记为(x_t, y_t)

（2）x_t是y_t的 Granger 原因，记为$(x_t \to y_t)$

（3）y_t是x_t的 Granger 原因，记为$(y_t \to x_t)$

（4）x_t和y_t互为因果，记为$(y_t \leftrightarrow x_t)$，这种情况也称为$x_t$和$y_t$之间存在反馈 Feedback 关系。

4.2 检验假设条件

依据 Granger 给出的定义，学者从不同角度构造检验统计量，给出了很多种 Granger 因果检验方法。这里介绍 Sargent 于 1976 年提出的直接 Granger 因果检验方法。

直接 Granger 因果检验认为绝大多数时间序列的生成过程是相互独立的，所以原假设是x_t不是y_t的 Granger 原因，备择假设是x_t是y_t的 Granger 原因，记作：

$$H_0: (x_t, y_t) \qquad H_1:\ (x_t \to y_t)$$

构造序列y_t的最优线性预测函数，记住：

$$y_t = \beta_0 + \sum_{k=1}^{p} \beta_k y_{t-k} + \sum_{k=1}^{q} \alpha_k x_{t-k} + \sum_{k=1}^{l} \gamma_k z_{t-k} + \varepsilon_t \qquad 10.19$$

式 10.19 中，p为序列y_t的自回归阶数；q为引入x_t序列的历史延迟阶数；z_t为其他自变量序列。原假设成立时，意味着$\alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$，所以假设条件可以等价表达为：

$$H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_q = 0 \qquad H_1: \alpha_1, \alpha_2, \ldots, \alpha_q \text{不全为零}$$

4.3 检验统计量

有多种方法构建 Granger 因果检验统计量，这里介绍 F 检验统计量的构造原理。在该检验方法下，需要拟合两个回归模型：

（1）在原假设成立的情况下，拟合序列y_t的有约束的预测模型（约束条件为$\alpha_1=\alpha_2=\cdots=\alpha_q=0$）为：

$$y_t=\beta_0+\sum_{k=1}^{p}\beta_k y_{t-k}+\sum_{k=1}^{l}\gamma_k z_{t-k}+\varepsilon_{1t} \qquad 10.20$$

对该模型进行方差分解：

$$SST=SSR_{yz}+SSE_1 \qquad 10.21$$

其中，$SST=\sum_{i=1}^{n}(y_i-\bar{y})^2$代表序列$y_t$的波动平方和，$n$ 为序列长度或观测值个数。SST可以分解成两个部分：一部分波动可以由 ,y_t 和 z_t 的历史信息$\{y_t,y_{t-1},y_{t-2},\dots,y_{t-p},z_t,z_{t-1},z_{t-2},\dots,z_{t-l}\}$解读，这部分记作$SSR_{yz}$；另一部分是不能由历史信息解读的，归为随机波动，记作有约束条件下的残差平方和SSE_1：

$$SSE_1=\sum_{t=1}^{n}{\varepsilon_{1t}}^2=SST-SSR_{yz} \qquad 10.22$$

（2）在备择假设成立的情况下，拟合序列y_t的无约束的预测模型：

$$y_t=\beta_0+\sum_{k=1}^{p}\beta_k y_{t-k}+\sum_{k=1}^{q}\alpha_k x_{t-k}+\sum_{k=1}^{l}\gamma_k z_{t-k}+\varepsilon_t \qquad 10.23$$

对该模型进行方差分解：

$$SST=SSR_{xyz}+SSE \qquad 10.24$$

其中，$SST=\sum_{i=1}^{n}(y_i-\bar{y})^2$代表序列$y_t$的波动平方和，$n$ 为序列长度或观测值个数。SST可以分解成两个部分：一部分波动可以由 x_t,y_t 和 z_t 的历史信息$\{x_t,x_{t-1},x_{t-2},\dots,x_{t-q},y_t,y_{t-1},y_{t-2},\dots,y_{t-p},z_t,z_{t-1},z_{t-2},\dots,z_{t-l}\}$解读，这部分记作$SSR_{xyz}$；实际上，还可以对$SSR_{xyz}$再进行分解，分解为$x_t$的影响和$y_t,z_t$的影响两个部分，即$SSR_{xyz}=SSR_{yz}+SSR_x$。剩下的不能由历史信息解读的，归为随机波动，记作无约束条件下的残差平方和SSE：

$$SSE=\sum_{t=1}^{n}{\varepsilon_t}^2=SST-SSR_{yz}-SSR_x \qquad 10.25$$

基于式 10.22 和式 10.25 构造 F 统计量：

$$F=\frac{(SSE_1-SSE)/q}{SSE/(n-q-p-1)}\sim F(q,n-q-p-1) \qquad 10.26$$

式 10.26 中，$SSE_1-SSE=SSR_x$，所以分子部分实际上是x_t的回归误差平方和除以它的自由度q，分母部分是无约束残差平方和除以它的自由度。SSR_x和SSE相互独立，所以它们

各自除以自由度服从 F 分布。若显著性水平为α，当 F 统计量大于$F_{1-\alpha}(q,n-p-q-1)$时，拒绝原假设，则x_t是y_t的 Granger 原因。

要注意，Granger 因果检验的结果严重依赖于解释变量的延迟阶数，即不同的延迟阶数 p 和 q 可能会得到不同的检验结果（通过影响SSE_1和SSE的值）。所以通常要借助y_t自相关系数图 PAC 确定p，借助互相关系数图 CCF 确定 q；或者多拟合几个不同延迟阶数的有约束模型和无约束模型，借助最小信息准则，使用 AIC/BIC/SBC 最小的无约束和有约束模型的残差平方和计算 F 统计量。

4.4 因果检验的其他问题

在做 Granger 因果检验时，要注意如下几个问题:

（1）检验结果只说明样本数据特征，如果换一批数据，或增加样本数据量，得出的因果判别可能会完全不一样。这也就是说，Granger 因果检验的结果会受到样本随机性的影响。样本容量越小，样本随机性的影响就越大。所以最好在样本容量比较大时进行 Granger 因果检验，以保证检验结果相对稳健。

（2）Granger 因果检验即使显著拒绝原假设，也不能说明两个序列间具有真正的因果关系。Granger 因果检验的构造思想是：使响应变量预测精度有显著提高的自变量可以视作响应变量的因。

这里面存在一个逻辑漏洞：如果变量 x 是变量 y 的因，那么知道 x 的信息对预测 y 是有帮助的，这个结论是对的。也就是说，因果性包含了预测精度的提高。但反过来，认为有助于预测精度提高的变量都是响应变量的因，就不一定正确了。比如说每天太阳快要升起的时候，公鸡都会打鸣。所以根据每天公鸡打鸣的时间，可以准确预测今天太阳升起的时间。根据 Granger 因果关系定义，可以认为公鸡打鸣是太阳升起的原因。显然这个因果结论是错误的。把公鸡杀了，太阳依然会升起。公鸡打鸣绝不是太阳升起的原因。这就说明由预测精度的提高反推因果性是不严谨的。

也就是说，**因果性可以推出预测精度提高，但预测精度提高不能等价推出因果性**。这就意味着，**在进行 Granger 因果检验时，即使得出因果关系显著成立的结论，也仅仅是预测精度提高的统计显著性判断，并不意味着两个变量之间一定存在真正的因果关系。**

Granger 因果检验是我们在处理复杂变量关系时使用的一个工具，Granger 因果检验的信息可以帮助我们思考模型的结构。它不一定百分百准确，但有它提供的信息比完全没有信息要强。