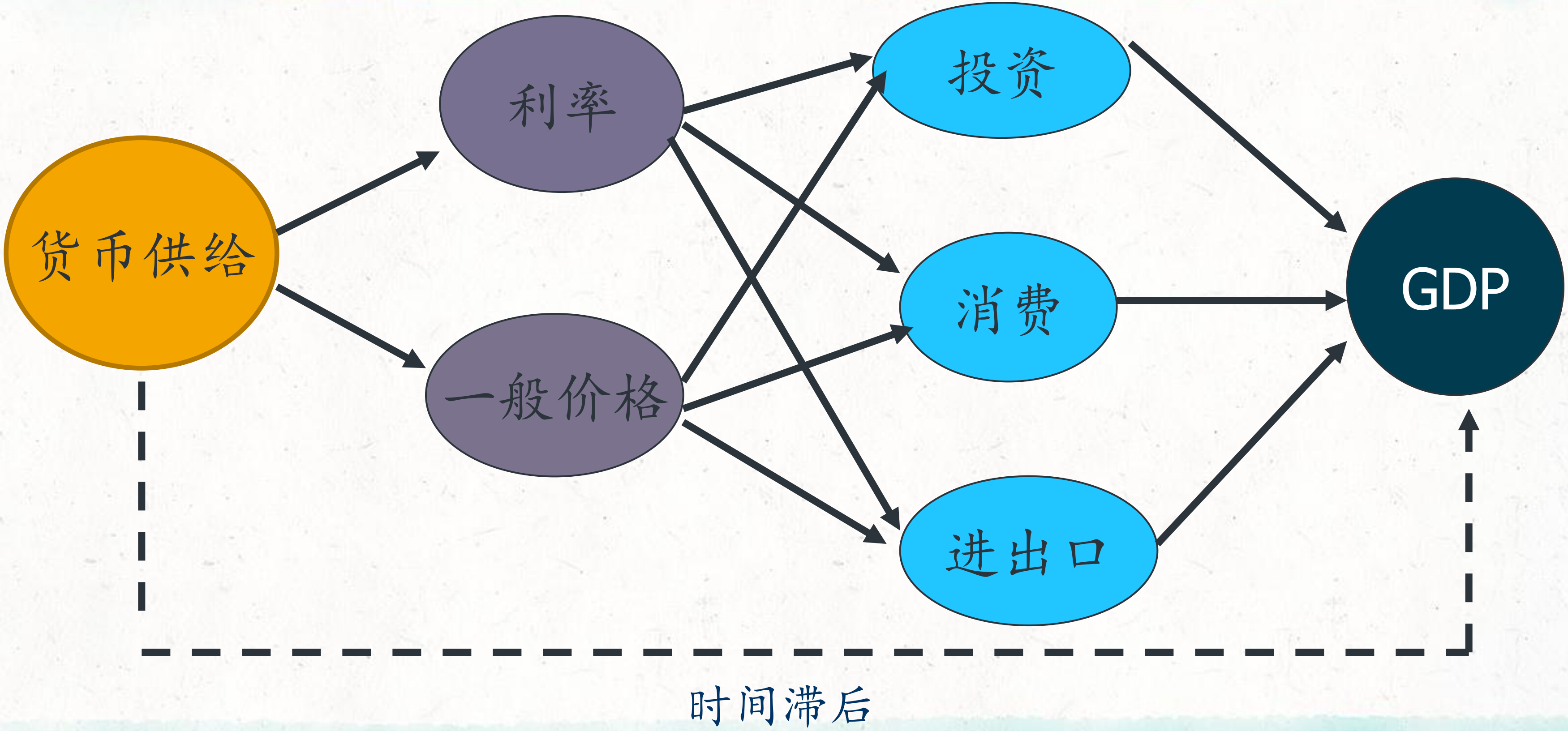
第七章 分布滞后模型与自回归模型

引子：货币政策效应的时滞



需要思考的问题:

在现实经济活动中，滞后现象是普遍存在的，这就要求我们在做经济分析时应该考虑时滞的影响。

解释变量与被解释变量的因果联系不可能在短时间内完成，在这一过程中通常都存在时间滞后，也就是说解释变量需要通过一段时间才能完全作用于被解释变量。

怎样才能把这类时间上滞后的经济关系纳入计量经济模型呢？

本 章 内 容:

- ◆滞后效应与滞后变量模型(分布滞后、自回归)
- ◆分布滞后模型的估计
- ◆自回归模型的构建
- ◆自回归模型的估计

此前讨论的模型变量间的关系是同时(瞬时、静态)的，实际不一定是这样。要反映不同时期变量之间的关系，需要引入滞后变量，使静态模型成为动态模型。

从时间关系上看：

变量间瞬时关系 (静态模型) ⟷ 不同时期变量间的关系 (动态模型)

第一节 滞后变量

一、滞后效应与滞后变量

滞后效应：

被解释变量受自身或其它变量过去值影响的现象，或被解释变量对解释变量的响应有一定的时间延滞，称为滞后效应

滞后值：

相对于某变量的本期值，该变量过去时期的数值称为滞后值

滞后变量：

模型中表示滞后值的变量称为滞后变量。

滞后变量分为：滞后解释变量，滞后被解释变量

二、滞后效应产生的原因

1、心理因素

心理习惯（惰性）：如收入增加后，消费习惯却有惯性心理预期：对未来的预期会影响本期的经济行为　如：现在收入增加——是否永久收入增加？预期价格会下降?

2、技术因素

如：◆投资⟶形成固定资产⟶经济增长　（有时滞）

◆货币供应量⟶通货膨胀（有时滞）

3、制度因素

契约与制度的改变有滞后，契约义务防碍对变化了的情况的决策

三、引入滞后变量的模型

1、 滞后变量引入模型的一般形式

可以引入滞后解释变量，也可以引入滞后被解释变量，最一般形式为

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_S X_{t-s} + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \cdots + \gamma_q Y_{t-q} + u_t$$

其中：α —截距项

β —解释变量及滞后值的参数

s —滞后解释变量的滞后期

γ —被解释变量滞后值的参数

q —滞后应变量的滞后期

2、分布滞后模型

- 模型中只含有滞后解释变量，被解释变量所受影响“分布在解释变量不同时期滞后值上”的模型。

一般形式：$Y_t=\alpha+\beta_0X_t+\beta_1X_{t-1}+\beta_2X_{t-2}+\cdots+\beta_sX_{t-s}+u_t$

或

$$Y_t=\alpha+\sum_{i=0}^{s}\beta_i\ X_{t-i}+u_t$$

（1）有限分布滞后模型：模型中解释变量滞后期的长度S是有限的，如S=K

$$Y_t=\alpha+\beta_0X_t+\beta_1X_{t-1}+\beta_2X_{t-2}+\cdots+\beta_kX_{t-k}+u_t$$

（2）无限分布滞后模型：模型中解释变量滞后期的长度是无限的，$s\longrightarrow\infty$

$$Y_t=\alpha+\beta_0X_t+\beta_1X_{t-1}+\beta_2X_{t-2}+\beta_3X_{t-3}+\cdots+u_t$$

分布滞后模型参数的经济意义：

模型 $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_s X_{t-s} + u_t$

短期乘数： β_0表示同期（滞后期为0）解释变量 X_t 变动一个单位，对本期因变量 Y_t 平均值的影响，称为短期乘数（即期乘数）

延迟乘数： β_1、$\beta_2\cdots$ 分别表示第 $t-1, t-2\cdots$ 时期的解释变量变动一个单位，对第t期因变量平均值的影响，分别称为延迟乘数或动态乘数。

长期乘数:经济处于稳定状态(长期平衡)时,所有变量为常量时, $\sum_{i=0}^{\infty}\beta_i = \beta_0 + \beta_1 + \beta_2\cdots$ 表示解释变量及其滞后值变动一个单位时，由于滞后效应对本期因变量 Y_t 平均值总的影响，称为长期乘数。

3、自回归模型

模型中的解释变量只包括解释变量的本期值和被解释变量若干期滞后值的模型。

一般形式：

$$Y_t = \alpha + \beta_0 X_t + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \cdots + \gamma_q Y_{t-q} + u_t$$

第二节 分布滞后模型及其估计

一、 分布滞后模型估计存在的问题

1、对于无限分布滞后模型：

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \cdots + u_t$$

●滞后项无限多 $\Longrightarrow$ 应估计的参数也无限多

●但样本观测值个数总是有限

—— 事实上不能直接估计其参数

2、对于有限分布滞后模型：

$$Y_t=\alpha+\beta_0X_t+\beta_1X_{t-1}+\beta_2X_{t-2}+\cdots+\beta_kX_{t-S}+u_t$$

可视为S+1个解释变量的模型去估计

但可能出现三个问题：

（1）解释变量滞后期长度如何确定（AIC/SIC，p164)

（2）滞后期较多，样本容量有限，自由度可能不够

（3）可能出现多重共线性：变量连续的逐期滞后值很可能高度相关

解决分布滞后模型估计问题的基本思路：

变换模型——设法把各滞后变量组合成为个数较少的新变量

目的：▸减少直接估计的参数个数 ▸增加自由度

▸避免多重共线性

二、有限分布滞后模型的估计方法

怎样变换模型？方法有多种

1、经验权数法

思想： 为减少要估计的参数个数，将各个解释变量组合为一个新变量，可对滞后变量的参数β作某种假定（施加某种约束），最简单的办法是对滞后变量指定一定的权数加以组合。权数的不同分布决定了滞后结构的不同类型

（1）递减滞后结构

假定： 解释变量对被解释变量的影响，随时间推移越来越小，按“近大远小”原则，X的权数由近到远逐步递减

例如：假定权数 W=1，1/2，1/4，1/8

加权的方法：

对于原模型

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_s X_{t-s} + u_t$$

●令新变量 $Z_t = w_0 X_t + w_1 X_{t-1} + \cdots + w_s X_{t-s}$

其中： $w_0 > w_1 > w_2 > \cdots > w_s$ 是预先指定的权数

例如，几何递减权数 $(\lambda < 1)$ $Z_t = \lambda^0 X_t + \lambda^1 X_{t-1} + \lambda^2 X_{t-2} + \cdots \lambda^s X_{t-s}$

如 $Z_t = X_t + \frac{1}{2} X_{t-1} + \frac{1}{4} X_{t-2} + \frac{1}{8} X_{t-3}$

●用 Z_t 代替各解释变量，模型变为：

$$Y_t = \alpha_0 + \alpha_1 Z_t + u_t$$

即 $Y_t = \alpha_0 + \alpha_1 w_0 X_t + \alpha_1 w_1 X_{t-1} + \cdots + \alpha_1 w_s X_{t-s} + u_t$

●用估计的 $\hat{\alpha}_1$ 可间接计算出各个 $\hat{\beta}_j (j = 1, 2, \cdots s)$

因为 $\hat{\beta}_0 = \hat{\alpha}_1 w_0, \hat{\beta}_1 = \hat{\alpha}_1 w_1, \cdots \hat{\beta}_s = \hat{\alpha}_1 w_s$

（2）不变滞后结构

假定：权数为常数，即

$$Z_t = wX_t + wX_{t-1} + \cdots + wX_{t-s} \qquad 或 \qquad Z_t = X_t + X_{t-1} + \cdots + X_{t-s}$$

例如： W＝1/4，1/4，1/4，1/4

（3）倒V形滞后结构

假定：滞后变量的权数先递增后递减，权数两头小中间大

如 $$Z_t = \frac{1}{4}X_t + \frac{1}{2}X_{t-1} + \frac{2}{3}X_{t-2} + \frac{1}{2}X_{t-3} + \frac{1}{4}X_{t-4}$$

权数: W = 1/4, 1/2, 2/3, 1/2, 1/4

经验权数法优缺点：

优点：简单易行，参数估计有一致性

缺点：滞后形式和权数指定有随意性

2、阿尔蒙法

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_s X_{t-s} + u_t$$

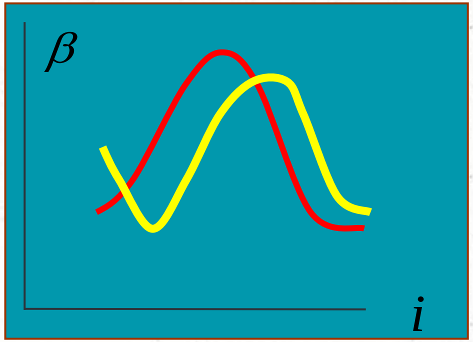
基本思想： 用某种多项式的方式减少待估参数的个数

前提： β_i随滞后期 i 而呈规律性变动，其变动可能呈某种曲线形式，

根据： 高等数学中“维尔斯特拉斯定理”：“一个有限闭区间的任何连续函数都可以用一个适当项的多项式去近似表示”。

如：$\beta_i = a_0 + a_1 i + a_2 i^2$ (A)

$\beta_i = a_0 + a_1 i + a_2 i^2 + a_3 i^3$ (B)



一般性: β_i 可以用滞后期 i 的m阶多项式去近似表示。

$$\beta_i = a_0 + a_1 i + a_2 i^2 + \cdots + a_m i^m$$

（滞后期 i =0,1,2,----s）

即

$$\beta_0 = a_0 + a_1 0 + a_2 0^2 + \cdots + a_m 0^m$$
$$\beta_1 = a_0 + a_1 1 + a_2 1^2 + \cdots + a_m 1^m$$
$$\beta_2 = a_0 + a_1 2 + a_2 2^2 + \cdots + a_m 2^m$$

$$\beta_s = a_0 + a_1 s + a_2 s^2 + \cdots + a_m s^m$$

关键:确定多项式的项次m: （经验方法：m至少比β和 i 的曲线的转向点个数大 1 即可）

原分布滞后模型

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_s X_{t-s} + u_t$$

$$Y_t = \alpha + \sum_{i=0}^{s} \beta_i X_{t-i} + u_t$$

将 $\beta_i = a_0 + a_1 i + a_2 i^2 + \cdots + a_m i^m$ 代入原模型得

$$Y_t = \alpha + \sum_{i=0}^{s} (a_0 + a_1 i + a_2 i^2 + \cdots + a_m i^m) X_{t-i} + u_t$$

或

$$Y_t = \alpha + a_0 \sum_{i=0}^{s} X_{t-i} + a_1 \sum_{i=0}^{s} i X_{t-i} + a_2 \sum_{i=0}^{s} i^2 X_{t-i} + \cdots + a_m \sum_{i=0}^{s} i^m X_{t-i} + u_t$$

注意：原模型有S+1个解释变量

变换后模型只有m+1个解释变量

整理后得

$$Y_t=\alpha+a_0\sum_{i=0}^{s}X_{t-i}+a_1\sum_{i=0}^{s}iX_{t-i}+a_2\sum_{i=0}^{s}i^2X_{t-i}+\cdots+a_m\sum_{i=0}^{s}i^mX_{t-i}+u_t$$

令 $\sum_{i=0}^{s}X_{t-i}=Z_{0t}$，$\sum_{i=0}^{s}iX_{t-i}=Z_{1t}$，$\sum_{i=0}^{s}i^2X_{t-i}=Z_{2t}$，$\sum_{i=0}^{s}i^mX_{t-i}=Z_{mt}$

即 $$Y_t=\alpha+a_0Z_{0t}+a_1Z_{1t}+a_2Z_{2t}+\cdots+a_mZ_{mt}+u_t$$

其中:：$Z_{0t},Z_{1t}\cdots Z_{mt}$ 是原滞后变量的线性组合 $(m<s)$

$$Z_{0t}=\sum_{i=0}^{s}i^0X_{t-i}=X_t+X_{t-1}+\cdots+X_{t-s}$$

$$Z_{1t}=\sum_{i=0}^{s}iX_{t-i}=1X_{t-1}+2X_{t-2}+\cdots+sX_{t-s}$$

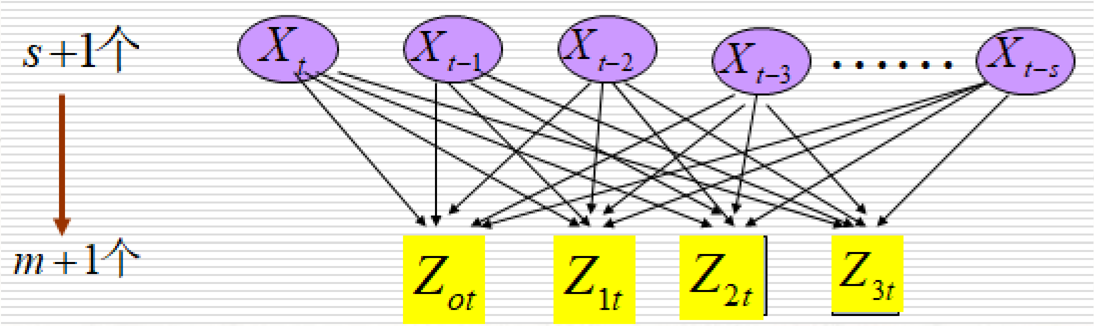
$$Z_{2t}=\sum_{i=0}^{s}i^2X_{t-i}=1X_{t-1}+4X_{t-2}+\cdots+s^2X_{t-s}$$

$$Z_{mt}=\sum_{i=0}^{s}i^mX_{t-i}=X_{t-1}+2^mX_{t-2}+\cdots+s^mX_{t-s}$$

以上过程中，滞后期数i为已知，只需估计出各个$\hat{a}$，即可计算出各个 $\hat{\beta}_i$ 因为 $\beta_i=a_0+a_1i+a_2i^2+\cdots+a_mi^m$

具体做法

- 设定多项式的项次m：一般取2—4即可，使m大大小于滞后期数s（只需估计较少参数，自由度得到保证，也减轻了多重共线性）
- 变换原滞后变量为 Z $\quad Z_{mt}=\sum_{i=0}^{s} i^{m} X_{t-i} \quad (m=0,1,\cdots m)$
- 用OLS法估计 $\hat{a}_j$
- 由 $\hat{a}_j$ 计算 $\hat{\beta}_j$



三、 无限分布滞后模型的估计——库伊克变换

基本思想：

无限分布滞后模型有无穷多个参数，无法直接估计。但可将无限分布滞后模型通过数学变换的方式转换为有限个参数的自回归模型，然后间接地估计其参数

前提条件：

- 所有的β_i 的符号都相同（即β_i不改变符号）
- β_i为几何递减滞后形式

$$\beta_i = \beta_0 \lambda^i \quad (0 \le \lambda \le 1)$$

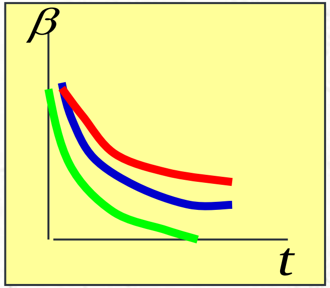
λ为分布滞后衰减率（近大远小）

具体作法：

原模型 $Y_t=\alpha+\beta_0 X_t+\beta_1 X_{t-1}+\beta_2 X_{t-2}+\beta_3 X_{t-3}+\cdots+u_t$

假定β_i为公比小于1的几何级数形式：

$$\beta_i=\beta_0\lambda^i \qquad (0\le\lambda\le 1)$$



代入原模型：

$$Y_t=\alpha+\beta_0 X_t+\beta_0\lambda X_{t-1}+\beta_0\lambda^2 X_{t-2}+\cdots+u_t \qquad (1)$$

（1）滞后一期并乘 λ：

$$\lambda Y_{t-1}=\lambda\alpha+\beta_0\lambda X_{t-1}+\beta_0\lambda^2 X_{t-2}+\beta_0\lambda^3 X_{t-3}+\cdots+\lambda u_{t-1} \qquad (2)$$

（1）式减（2）式：

$$Y_t-\lambda Y_{t-1}=\alpha(1-\lambda)+\beta_0 X_t+(u_t-\lambda u_{t-1})$$

(上页模型）

$$Y_t - \lambda Y_{t-1} = \alpha(1-\lambda) + \beta_0 X_t + (u_t - \lambda u_{t-1})$$

移项 $$Y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda Y_{t-1} + (u_t - \lambda u_{t-1})$$

令 $$\alpha^* = \alpha(1-\lambda) \qquad \beta_0^* = \beta_0$$

$$\beta_1^* = \lambda \qquad v_t = u_t - \lambda u_{t-1}$$

得 $$Y_t = \alpha^* + \beta_0^* X_t + \beta_1^* Y_{t-1} + v_t$$

这样，将无限分布滞后模型巧妙地变换为了一阶自回归模型，若估计出 $\alpha^*, \beta_0^* = \beta_0, \beta_1^* = \lambda$ 等，可计算出原模型各个参数的估计值 $\beta_i = \beta_0 \lambda^i$

库伊克变换的优点：

- 将有无穷多个参数要估计的无限分布滞后模型，变换为只有三个参数的自回归模型，使参数估计变为可能。
- 极大地减少了自由度的损失 。
- 解决了滞后长度难以确定的问题。
- 用被解释变量滞后值取代大量滞后解释变量，从而消除了多重共线性。

库伊克变换存在的问题：

- 有严格的假定条件（按固定比例递减），不一定符合经济问题的实际。
- 把随机变量 Y_{t-1} 引入了解释变量，不一定符合基本假定。
- 随机扰动 $v_t = u_t - \lambda u_{t-1}$ 可能自相关。
- 只是纯粹的数学运算的结果，缺乏经济理论依据。

第三节 自回归模型的构建

问题的提出：

库伊克变换形式巧妙，缺乏建模的经济背景。但是也可以从经济问题出发得到类似的模型形式，说明库伊克变换的经济背景

一、自适应预期模型

根据预期理论：人们的经济行为不仅受当前经济因素影响，而且受人们对某些经济变量未来走势的“预期”的影响，因此可以将某些变量的预期值作为解释变量

$$Y_t = \beta_0 + \beta_1 X_t^* + u_t$$

其中：X_t^* 是对变量X的预期水平

例如：货币需求Y是预期利率 X_t^* 的函数

问题:预期变量的预期值是不可观测的，只能根据预期形成机理对它作出某种假定

自适应预期理论（一种预期形成机理的假定）

为了作出合理的预期，可以根据过去所作预期的经验，不断修正当前的预期。按过去预期值与实际值偏差的一定比例去修正其预期值

本期预期值=上期预期值 + 修正值

$$X_t^* = X_{t-1}^* + \gamma(X_t - X_{t-1}^*)$$

其中修正值是上期预期误差的一部分,γ 是修正系数

或改写为：$$X_t^* = \gamma X_t + (1-\gamma)X_{t-1}^*$$

预期形成机理：说明本期预期值 是本期实际值与上期预期值的加权平均，权数是γ 和（1- γ）

注意理解: 第 t-1期作的预期 X_{t-1}^* 是对第 t期的 X_t 作的预期

建模代换：

思想：因预期值无法观测，设法通过代换在模型中避开直接使用预期值

作法：

●将建立在自适应预期机理基础上的预期值代入原模型

原模型 $$Y_t = \beta_0 + \beta_1 X_t^* + u_t$$

得 $$Y_t = \beta_0 + \beta_1[\gamma X_t + (1-\gamma) X_{t-1}^*] + u_t$$

即 $$Y_t = \beta_0 + \beta_1 \gamma X_t + \beta_1(1-\gamma) X_{t-1}^* + u_t \qquad \text{(A)}$$

●将原模型滞后一期并乘（1-γ ）得

$$(1-\gamma)Y_{t-1} = \beta_0(1-\gamma) + \beta_1(1-\gamma)X_{t-1}^* + (1-\gamma)u_{t-1} \qquad \text{（B）}$$

$$Y_t = \beta_0 + \beta_1 \gamma X_t + \beta_1(1-\gamma)X^*_{t-1} + u_t \quad （A）$$

$$(1-\gamma)Y_{t-1} = \beta_0 - \beta_0\gamma + \beta_1(1-\gamma)X^*_{t-1} + (1-\gamma)u_{t-1} \quad （B）$$

●以上两式相减 [（A）—（B）] 得:

$$Y_t - (1-\gamma)Y_{t-1} = \gamma\beta_0 + \gamma\beta_1 X_t + [u_t - (1-\gamma)u_{t-1}]$$

移项 $$Y_t = \underline{\gamma\beta_0} + \underline{\gamma\beta_1} X_t + \underline{(1-\gamma)}Y_{t-1} + \underline{[u_t - (1-\gamma)u_{t-1}]}$$

令 $\beta_0^* = \gamma\beta_0$，$\beta_1^* = \gamma\beta_1$，$\beta_2^* = 1-\gamma$，$v_t = u_t - (1-\gamma)u_{t-1}$

则 $$Y_t = \beta_0^* + \beta_1^* X_t + \beta_2^* Y_{t-1} + v_t$$

这是一个与库伊克变换相似的一阶自回归模型， 通过 $\hat{\beta}_j^*$ 可以计算出自适应预期模型的参数

二、局部调整模型

基本思想：

● 在经济管理中，常需要研究最适合的预期水平。

例如：预期的最佳货币供应量（相对于某经济发展水平）

预期的最佳商品储备（相对于某销售量）

预期最佳的资本存量（相对于某产出量）

这时需要将预期值作为被解释变量，将某些现期值作为解释变量。

●例如预期的最适宜资本存量水平 Y_t^* 可能与产出X有关，可建立模型：

$$Y_t^* = \beta_0 + \beta_1 X_t + u_t \quad (1)$$

存在的问题：Y_t^* 不能直接观测

资本投资理论的存货局部调整原理：

企业总要调整其资本存量Y_t，使其逐步接近预期的最适宜水平Y_t^*，由于种种限制这种调整只能逐步进行，认为实际的调整量只是预期调整量的一部分，假定调整机理为局部调整模型：

$$Y_t - Y_{t-1} = \delta(Y_t^* - Y_{t-1})$$

其中：$Y_t - Y_{t-1}$ 为实际调整量

$Y_t^* - Y_{t-1}$为预期的最适宜调整量

δ为调整系数 $0 \leqslant \delta \leqslant 1$

也可表示为 $Y_t = \delta Y_t^* + (1-\delta)Y_{t-1}$（2）

可见，Y_t 是Y_t^*和Y_{t-1}的加权平均数，权数为δ和 $1-\delta$

●然而，预期变量 Y_t^*不能观测，为代换Y_t^*，将（1）式代入（2）式：

$$Y_t = \delta(\beta_0 + \beta_1 X_t + u_t) + (1-\delta)Y_{t-1}$$

或

$$Y_t = \delta\beta_0 + \delta\beta_1 X_t + (1-\delta)Y_{t-1} + \delta u_t$$

令 $\beta_0^* = \delta\beta_0 \quad \beta_1^* = \delta\beta_1 \quad \beta_2^* = 1-\delta \quad v_t = \delta u_t$

得

$$Y_t = \beta_0^* + \beta_1^* X_t + \beta_2^* Y_{t-1} + v_t$$

这是由投资理论导出的一阶自回归模型

特点： $v_t = \delta u_t$ 较简单，且不导致自相关

结论：对比三种自回归模型的异同

共同点：模型最终形式都是一阶自回归模型

模型	模 型 形 式（自回归模型）	建模思想和依据	随机误差项结构和性质
库伊克	$Y_t=\alpha(1-\lambda)+\beta_0X_t+\lambda Y_{t-1}+(u_t-\lambda u_{t-1})$	参数几何级数递减数学变换	可能导致自相关 $v_t=u_t-\lambda u_{t-1}$
自适应预期	$Y_t=\gamma\beta_0+\gamma\beta_1X_t+(1-\gamma)Y_{t-1}+[u_t-(1-\gamma)u_{t-1}]$	自适应预期假定	可能导致自相关 $v_t=u_t-(1-\gamma)u_{t-1}$
局部调整	$Y_t=\delta\beta_0+\delta\beta_1X_t+(1-\delta)Y_{t-1}+\delta u_t$	局部调整机理	不导致自相关 $v_t=\delta u_t$

第四节 自回归模型的估计

一、自回归模型估计存在的问题

对一阶自回归模型 $Y_t = \alpha + \beta_0 X_t + \beta_1 Y_{t-1} + u_t$

存在问题：

- 出现了随机解释变量Y_{t-1}，且Y_{t-1}可能与u_t相关
- u_t 可能自相关：可以证明库伊克模型和自适应预期模型的随机扰动项都会导致自相关，只有局部调整模型的随机扰动无自相关

后果：违反基本假定，OLS估计不仅是有偏的，而且在大样本时是不一致的(证明较复杂，略)

要解决的问题：

- 设法消除Y_{t-1}与u_t 的相关性（寻求方法）。
- 检验u_t是否存在自相关（寻求自相关的检验方式）

二、工具变量法（消除 Y_{t-1} 与 u_t 的相关性）

基本思想

在模型 $Y_t = \alpha + \beta X_t + u_t$ 中，若是 X_t 与 u_t 相关，将违反基本假定。但如果能找到另一个变量 Z_t，使 Z_t 与 X_t 高度相关，但与 u_t 不相关，则可用 Z_t 代替 X_t 进行回归。这样的变量 Z_t 称为工具变量。

可以证明用工具变量法估计的参数是一致估计。

工具变量的选择条件：

- 与所替代的解释变量高度相关
- 与随机扰动项不相关
- 与模型中其他解释变量不相关（避免多重共线性）

（只帮忙不添乱）

具体作法：如何选择 Y_{t-1} 的工具变量？

1、用 $\hat{Y}_{t-1}$ 作工具变量代替 Y_{t-1}（为什么可这样选？）

●将 Y_t 对X滞后值回归

$$Y_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_s X_{t-s} + u_t$$

(滞后期 s 一般可选2、3）

●估计出参数后，滞后一期计算 $\hat{Y}_{t-1}$

$$\hat{Y}_{t-1} = \hat{\beta}_0 + \hat{\beta}_1 X_{t-2} + \hat{\beta}_2 X_{t-3} + \cdots + \hat{\beta}_s X_{t-s-1}$$

●用 $\hat{Y}_{t-1}$ 作工具变量代替 Y_{t-1}

$$Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 \hat{Y}_{t-1} + v_t$$

效果：小样本时有偏，大样本时渐近一致

2、用 X_{t-1} 作工具变量代替 Y_{t-1}

为什么可这样选？通常 X_{t-1} 与 Y_{t-1} 相关，与 v_t 不相关

问题：

- ► X_{t-1} 与 X_t 可能发生多重共线性
- ► v_t 仍然可能自相关

解决的办法：

- ►检验多重共线性是否严重
- ►检验 v_t 是否自相关

也可以找其他的符合条件的工具变量

三、自回归模型中自相关的检测——德宾 h 检验

目的： 检验自回归模型中是否存在自相关，分析估计结果的合理性

存在的问题：

- **回顾：** 检验自相关的 d 统计量有检验条件（P142）

  解释变量全是非随机变量；

  解释变量中没有滞后内生变量，即没有被解释变量滞后值（不是自回归）

- d 统计量检验不适于自回归模型，因为此时d总是趋近于2，会存在阻碍发现自相关的“内生偏倚”

解决的办法:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 Y_{t-1} + u_t$$

德宾提出 h 统计量，可检验自回归模型中针对 $H_0: \rho = 0$ 的自相关

$$h = \hat{\rho}\sqrt{\frac{n}{1-n[Var(\hat{\beta}_1)]}}$$

其中： n ——有效样本容量

$Var(\hat{\beta}_1)$——滞后应变量 Y_{t-1}的参数的方差

$\hat{\rho}$——一阶自相关系数的估计值 $\hat{\rho} = \dfrac{\sum e_t e_{t-1}}{\sum e_t^2}$

已知d时也可用 $\hat{\rho} \approx 1 - \dfrac{d}{2}$ 近似计算

大样本时h服从标准正态分布，可用于检验是否存在自相关

具体作法：

- 对一阶自回归模型 $Y_t = \alpha + \beta_0 X_t + \beta_1 Y_{t-1} + u_t$

  直接用OLS法估计其参数，并得 $Var(\hat{\beta}_1)$ 和 d 统计量

- 用 $Var(\hat{\beta}_1)$、d 统计量、n等数据计算 h 统计量

- 对于 $H_0: \rho = 0$

  给定显著性水平 α，查标准正态分布表得临界值 $z_{\alpha/2}$

  若 $|h| \geq z_{\alpha/2}$,拒绝 $H_0: \rho = 0$，存在自相关

  若 $|h| < z_{\alpha/2}$,不拒绝 $H_0: \rho = 0$，不存在自相关

注意：

- h 检验与模型中有多少个X 变量无关，计算h只考虑 Y_{t-1} 系数的方差（适用于任意阶自回归的原模型）
- h 检验只适用于大样本，小样本时效果差

第五节 案例

案例一：中国货币供给对物价变动影响滞后性的研究

（一）问题提出：货币供应量对物价的影响存在一定时滞。西方国家的通货膨胀时滞大约为2—3个季度。在中国货币供给的变化对物价也具有滞后影响，但滞后期究竟有多长？

（二）模型设定：为了考察货币供应量的变化对物价的影响，我们用广义货币M2的月增长量M2Z作为解释变量，以居民消费价格月度同比指数TBZS为被解释变量进行研究。首先建立如下回归模型

$$TBZS_t = \alpha + \beta_0 M2Z_t + u_t$$

（三）收集数据：采集1996－2008年全国广义货币供应量和物价指数的月度数据（表7.2，P177-179）

1、当期数据回归结果

Dependent Variable: TBZS
Method: Least Squares
Date: 12/09/19 Time: 12:52
Sample (adjusted): 1996M02 2008M11
Included observations: 154 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	101.3693	0.347947	291.3353	0.0000
M2Z	0.295873	0.099444	2.975258	0.0034

R-squared	0.055033	Mean dependent var	102.1383
Adjusted R-squared	0.048816	S.D. dependent var	2.964169
S.E. of regression	2.890915	Akaike info criterion	4.973925
Sum squared resid	1270.323	Schwarz criterion	5.013366
Log likelihood	-380.9922	Hannan-Quinn criter.	4.989946
F-statistic	8.852160	Durbin-Watson stat	0.144830
Prob(F-statistic)	0.003406		

X的t统计量值显著，表明当期货币供应量的变化对当期物价水平的影响在统计意义上有一定影响，但没有显现出这种影响的滞后性。

2、滞后6个月的分布滞后模型

分析货币供应量变化影响物价的滞后性，作滞后6个月的分布滞后模型的估计，在Eviews工作文档的“Equation Spesification”方程设定窗口中，输入:“TBZS C M2Z M2Z(-1) M2Z(-2) M2Z(-3) M2Z(-4) M2Z(-5) M2Z(-6)” （也可以输入M2Z(-1 to -6)，下同）

结果显示(见下页):

- M2Z各滞后期的系数逐步增加，表明当期货币供应量的变化对物价水平的影响要经过一段时间才能逐步显现。
- 但各滞后期的系数的t统计量值有些不显著，因此还不能据此判断滞后期究竟有多长。

View | Proc | Object | | Print | Name | Freeze | | Estimate | Forecast | Stats | Resids

Dependent Variable: TBZS
Method: Least Squares
Date: 12/09/19 Time: 12:59
Sample (adjusted): 1996M08 2008M11
Included observations: 148 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	99.22661	0.386141	256.9702	0.0000
M2Z	0.047766	0.096426	0.495367	0.6211
M2Z(-1)	0.134020	0.091565	1.463668	0.1455
M2Z(-2)	0.157368	0.090457	1.739691	0.0841
M2Z(-3)	0.152117	0.092776	1.639618	0.1033
M2Z(-4)	0.179926	0.090157	1.995700	0.0479
M2Z(-5)	0.166697	0.092253	1.806941	0.0729
M2Z(-6)	0.179974	0.097170	1.852159	0.0661

R-squared	0.305349	Mean dependent var	101.8561
Adjusted R-squared	0.270616	S.D. dependent var	2.659733
S.E. of regression	2.271517	Akaike info criterion	4.531312
Sum squared resid	722.3708	Schwarz criterion	4.693323
Log likelihood	-327.3171	Hannan-Quinn criter.	4.597136
F-statistic	8.791432	Durbin-Watson stat	0.095997
Prob(F-statistic)	0.000000		

3、滞后12个月的分布滞后模型的估计

Dependent Variable: TBZS
Method: Least Squares
Date: 12/09/19 Time: 13:00
Sample (adjusted): 1997M02 2008M11
Included observations: 142 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	98.19975	0.313325	313.4122	0.0000
M2Z	-0.064923	0.086361	-0.751762	0.4536
M2Z(-1)	0.079507	0.078461	1.013327	0.3128
M2Z(-2)	0.068629	0.081667	0.840355	0.4023
M2Z(-3)	0.099556	0.082280	1.209969	0.2285
M2Z(-4)	0.132429	0.082881	1.597825	0.1125
M2Z(-5)	0.044290	0.082215	0.538714	0.5910
M2Z(-6)	0.067894	0.082124	0.826729	0.4099
M2Z(-7)	0.131624	0.082236	1.600565	0.1119
M2Z(-8)	0.152601	0.082487	1.849993	0.0666
M2Z(-9)	0.085495	0.082246	1.039503	0.3005
M2Z(-10)	0.078295	0.081444	0.961335	0.3382
M2Z(-11)	0.204747	0.094826	2.159175	0.0327
M2Z(-12)	0.288988	0.100707	2.869585	0.0048

R-squared	0.554030	Mean dependent var	101.6366
Adjusted R-squared	0.508737	S.D. dependent var	2.482034
S.E. of regression	1.739662	Akaike info criterion	4.038645
Sum squared resid	387.3823	Schwarz criterion	4.330065
Log likelihood	-272.7438	Hannan-Quinn criter.	4.157066
F-statistic	12.23193	Durbin-Watson stat	0.201552
Prob(F-statistic)	0.000000		

(接上页)

从X到X(-10)回归系数都不显著异于零，而X(-11)的回归系数t统计量值为2.50，在5%显著性水平下拒绝系数为零的原假设。

这一结果表明，当期货币供应量变化对物价水平的影响在经过11个月后明显地显现出来。

为了考察货币供应量变化对物价水平影响的持续期，再作滞后18个月的分布滞后模型的估计

4、滞后18个月的分布滞后模型

Dependent Variable: TBZS
Method: Least Squares
Date: 12/09/19 Time: 13:04
Sample (adjusted): 1997M08 2008M11
Included observations: 136 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	97.59653	0.286256	340.9417	0.0000
M2Z	-0.019521	0.077185	-0.252917	0.8008
M2Z(-1)	0.015063	0.077286	0.194904	0.8458
M2Z(-2)	-0.020540	0.079295	-0.259027	0.7961
M2Z(-3)	0.004308	0.079056	0.054496	0.9566
M2Z(-4)	0.001522	0.081215	0.018744	0.9851
M2Z(-5)	0.004786	0.082489	0.058024	0.9538
M2Z(-6)	-0.011762	0.081670	-0.144018	0.8857
M2Z(-7)	0.066961	0.078720	0.850623	0.3967
M2Z(-8)	0.091757	0.078392	1.170487	0.2442
M2Z(-9)	0.043119	0.078385	0.550092	0.5833
M2Z(-10)	0.036499	0.077371	0.471739	0.6380
M2Z(-11)	0.164543	0.087029	1.890671	0.0612
M2Z(-12)	0.214225	0.094830	2.259038	0.0257
M2Z(-13)	0.231705	0.094485	2.452292	0.0157
M2Z(-14)	0.212450	0.095659	2.220911	0.0283
M2Z(-15)	0.215433	0.097011	2.220696	0.0283
M2Z(-16)	0.172157	0.096130	1.790880	0.0759
M2Z(-17)	0.109469	0.096874	1.130012	0.2608
M2Z(-18)	0.114871	0.092097	1.247289	0.2148

R-squared	0.685178	Mean dependent var	101.5537
Adjusted R-squared	0.633612	S.D. dependent var	2.494614
S.E. of regression	1.509988	Akaike info criterion	3.797134
Sum squared resid	264.4876	Schwarz criterion	4.225466
Log likelihood	-238.2051	Hannan-Quinn criter.	3.971197
F-statistic	13.28749	Durbin-Watson stat	0.197989
Prob(F-statistic)	0.000000		

滞后18个月的分布滞后模型回归结果分析

分析什么？参数的变动规律及统计检验结果

回归系数的变动规律:

从滞后12个月开始，货币供应量变化对物价水平的影响明显增加，在滞后14个月时达到最大，然后逐步下降。

参数的检验结果:

滞后12个月开始t统计量值显著，一直到滞后16个月为止，从滞后第17个月开始t值变得不显著。

判断：在中国，货币供应量变化对物价水平的影响具有明显的滞后性，滞后期大约为四个季度，而且滞后影响具有持续性，持续的长度大约为半年，其影响力度先递增然后递减，滞后结构为 Λ 型（倒V型）。

回归结果显示：回归方程的可决系数不高，DW值也偏低，表明除了货币供应量外，还有其他因素影响物价变化；同时，过多的滞后变量也可能引起多重共线性问题。

但是如果我们分析的重点只是货币供应量变化对物价影响的滞后性，上述结果已能说明问题。

5、用自回归模型代替分布滞后模型

估计如下自回归模型

$$TBZS_t = \alpha + \beta_0 M2Z_t + \gamma TBZS_{t-1} + u_t$$

Dependent Variable: TBZS
Method: Least Squares
Date: 12/09/19 Time: 13:14
Sample (adjusted): 1996M02 2008M11
Included observations: 154 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.446892	1.675727	2.653709	0.0088
M2Z	0.045873	0.021161	2.167762	0.0317
TBZS(-1)	0.954894	0.016494	57.89322	0.0000

R-squared	0.959262	Mean dependent var	102.1383
Adjusted R-squared	0.958722	S.D. dependent var	2.964169
S.E. of regression	0.602227	Akaike info criterion	1.842924
Sum squared resid	54.76429	Schwarz criterion	1.902085
Log likelihood	-138.9051	Hannan-Quinn criter.	1.866955
F-statistic	1777.805	Durbin-Watson stat	1.665626
Prob(F-statistic)	0.000000		

案例二:某地制造业库存量与销售额的关系

模型: $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \mu_t$

样本数据:

年份	销售额X	库存量Y	年份	销售额X	库存量Y
1987	26.48	45.069	1997	41.003	68.221
1988	27.74	50.642	1998	44.869	77.965
1989	28.236	51.871	1999	46.449	84.655
1990	27.28	52.07	2000	50.282	90.815
1991	30.219	52.709	2001	53.555	97.074
1992	30.796	53.814	2002	52.859	101.64
1993	30.896	54.939	2003	55.917	102.44
1994	33.113	58.123	2004	62.017	107.71
1995	35.023	60.043	2005	71.398	120.87
1996	37.335	63.383	2006	82.078	147.13

分布滞后模型与自回归模型的建立

估计分布滞后模型:键入命令LS Y C X X(-1) X(-2) X(-3)/OK

得到原模型的估计结果

判断:

原分布滞后模型明显存在多重共线性。（注意：样本区间的变化）

Dependent Variable: Y
Method: Least Squares
Date: 10/15/07   Time: 07:14
Sample(adjusted): 1990 2006
Included observations: 17 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6.503312	2.236561	-2.907728	0.0131
X	0.587778	0.250807	2.343551	0.0371
X(-1)	1.261146	0.460984	2.735769	0.0181
X(-2)	0.649926	0.482753	1.346290	0.2031
X(-3)	-0.500561	0.343335	-1.457938	0.1705

R-squared	0.996257	Mean dependent var	81.97653
Adjusted R-squared	0.995009	S.D. dependent var	27.85539
S.E. of regression	1.967864	Akaike info criterion	4.431703
Sum squared resid	46.46986	Schwarz criterion	4.676766
Log likelihood	-32.66948	F-statistic	798.4718
Durbin-Watson stat	1.500070	Prob(F-statistic)	0.000000

1、经验加权法

取权数为： $w_0 = 1, w_1 = 1/2, w_2 = 1/4, w_3 = 1/8$

在“Workfile”表中点“Genr”，在出现的对话框中输入：

Z1=X+X(-1)/2+X(-2)/4+X(-3)/8

点“OK”即生成Z1，

作回归：

“LS Y C Z1”/OK，

即得回归结果。

注意因生成滞后变量减少了3个样本，这时样本区间为：

1990-2006

Dependent Variable: Y
Method: Least Squares
Date: 10/15/07 Time: 07:18
Sample(adjusted): 1990 2006
Included observations: 17 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.950580	1.930582	-3.082272	0.0076
Z1	1.064556	0.022349	47.63335	0.0000

R-squared	0.993432	Mean dependent var	81.97653
Adjusted R-squared	0.992995	S.D. dependent var	27.85539
S.E. of regression	2.331456	Akaike info criterion	4.640994
Sum squared resid	81.53530	Schwarz criterion	4.739019
Log likelihood	-37.44845	F-statistic	2268.936
Durbin-Watson stat	1.350648	Prob(F-statistic)	0.000000

根据经验加权的结果计算原模型参数估计值

原模型： $Y_t=\alpha+\beta_0X_t+\beta_1X_{t-1}+\beta_2X_{t-2}+\beta_3X_{t-3}+\mu_t$

经验加权模型： $Y_t=\alpha_0+\alpha_1Z_t+u_t$

或 $Y_t=\alpha_0+\alpha_1w_0X_t+\alpha_1w_1X_{t-1}+\alpha_1w_2X_{t-2}+\alpha_1w_3X_{t-3}+u_t$

经验加权估计结果：

$\hat{\alpha}=\hat{\alpha}_0=-5.9506$ $\quad\hat{\beta}_0=1*1.064556$ $\quad\hat{\beta}_1=\frac{1}{2}*1.064556=0.532278$

$\hat{\beta}_2=\frac{1}{4}*1.064556=0.266139$ $\quad\hat{\beta}_3=\frac{1}{8}*1.064556=0.1330695$

得出原分布滞后模型的估计结果：

$$\hat{Y}_t=-5.9536+1.0646X_t+0.5323X_{t-1}+0.2661X_{t-2}+0.1331X_{t-3}$$

2、阿尔蒙法

方法1:生成新变量法

有限分布滞后模型：$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + u_t$

将系数 β_i (i=0,1,2,3)用二次多项式近似，原模型可变为

$$Y_t = \alpha + \alpha_0 Z_{0t} + \alpha_1 Z_{1t} + \alpha_2 Z_{2t} + u_t$$

点击“Genr”工具栏，生成新变量

$$Z_{0t} = X + X(-1) + X(-2) + X(-3)$$

$$Z_{1t} = \sum_{i=0}^{s} iX_{t-i} = 0*X + 1*X(-1) + 2*X(-2) + 3*X(-3)$$

$$Z_{2t} = \sum_{i=0}^{s} i^2 X_{t-i} = 1*X(-1) + 4*X(-2) + 9X(-3)$$

Dependent Variable: Y
Method: Least Squares
Date: 10/15/07 Time: 08:07
Sample(adjusted): 1990 2006
Included observations: 17 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6.415558	2.128045	-3.014767	0.0100
Z0T	0.630109	0.179003	3.520096	0.0038
Z1T	0.988165	0.524876	1.882665	0.0823
Z2T	-0.461110	0.181050	-2.546867	0.0243

R-squared	0.996237	Mean dependent var	81.97653
Adjusted R-squared	0.995369	S.D. dependent var	27.85539
S.E. of regression	1.895631	Akaike info criterion	4.319305
Sum squared resid	46.71442	Schwarz criterion	4.515355
Log likelihood	-32.71409	F-statistic	1147.287
Durbin-Watson stat	1.514374	Prob(F-statistic)	0.000000

作回归“Y C Z0t Z1t Z2t”/OK，得回归结果

$\hat{\alpha} = -6.4156$ $\hat{a}_0 = 0.6301$

$\hat{a}_1 = 0.9882$ $\hat{a}_2 = -0.4611$

由 $\hat{\alpha}=-6.4156 \quad \hat{a}_0=0.6301$

$\hat{a}_1=0.9882 \quad \hat{a}_2=-0.4611$

根据阿尔蒙变换公式求出原分布滞后模型的各个参数，

$$\hat{\beta}_0=\hat{\alpha}_0=0.6301$$

$$\hat{\beta}_1=\hat{\alpha}_0+1\hat{\alpha}_1+1^2\hat{\alpha}_2=0.6301+0.9882+(-0.4611)=1.1572$$

$$\hat{\beta}_2=\hat{\alpha}_0+2\hat{\alpha}_1+2^2\hat{\alpha}_2=0.6301+2\times0.9882+4\times(-0.4611)=0.7621$$

$$\hat{\beta}_3=\hat{\alpha}_0+3\hat{\alpha}_1+3^2\hat{\alpha}_2=0.6301+3\times0.9882+9\times(-0.4611)=\text{-}0.5552$$

估计结果为

$$\hat{Y}_t=-6.4156+0.6301X_t+1.1572X_{t-1}+0.7621X_{t-2}-0.5552X_{t-3}$$

对比用“PDL”命令估计的分布滞后模型结果 (见后面)

$$\hat{Y}_t=-6.4156+0.6301X_t+1.1572X_{t-1}+0.7620X_{t-2}-0.5554X_{t-3}$$

方法2:Eviews指令“PDL”用于估计分布滞后模型

在Eviews中键入命令：“**ls Y C PDL(X，3，2)**”

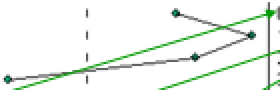
其中，“**PDL指令**”表示进行多项式分布滞后（Polynomial Distributed Lags）模型的估计，括号中的**3**表示**X**的分布滞后长度，**2**表示多项式的阶数。

在**Estimation Settings**栏中选择**Least Squares(**最小二乘法**)**，点击**OK**，得回归分析结果

Dependent Variable: Y
Method: Least Squares
Date: 10/15/07 Time: 08:39
Sample(adjusted): 1990 2006
Included observations: 17 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6.415558	2.128045	-3.014767	0.0100
PDL01	1.157164	0.195764	5.911024	0.0001
PDL02	0.065945	0.175906	0.374888	0.7138
PDL03	-0.461110	0.181050	-2.546867	0.0243

R-squared	0.996237	Mean dependent var	81.97653
Adjusted R-squared	0.995369	S.D. dependent var	27.85539
S.E. of regression	1.895631	Akaike info criterion	4.319305
Sum squared resid	46.71442	Schwarz criterion	4.515355
Log likelihood	-32.71409	F-statistic	1147.287
Durbin-Watson stat	1.514374	Prob(F-statistic)	0.000000



Lag Distribution of X	i	Coefficient	Std. Error	T-Statistic
	0	0.63011	0.17900	3.52010
	1	1.15716	0.19576	5.91102
	2	0.76200	0.17805	4.27971
	3	-0.55539	0.25538	-2.17474
Sum of Lags		1.99389	0.06778	29.4162

$$\hat{Y}_t = -6.4156 + 0.6301X_t + 1.1572X_{t-1} + 0.7620X_{t-2} - 0.5554X_{t-3}$$

需要指出的是，用“PDL”估计分布滞后模型时，Eviews所采用的滞后系数多项式变换不是形如 $\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2$ 的阿尔蒙多项式，而是阿尔蒙多项式的派生形式：

$$\beta_i = \alpha_0 + \alpha_1 (i-1) + \alpha_2 (i-1)^2$$

因此，输出结果中估计的PDL01、PDL02、PDL03系数与用阿尔蒙多项式分步估计的系数不一致。但同前面分步计算的结果相比，最终的分布滞后系数估计结果相同。底部的折线图是分布滞后模型解释变量参数估计值的线性图，也就是β_0、β_1、β_2、β_3估计值构成的折线图。该图形需要逆时针旋转90度来看，图中虚线表示0值线。