第七章： GLS( generalized least squares )估计及相关话题

如果系数要通过OLS方法有效的进行估计，误差项必须没有相关性且同方差，这也是高斯-马尔科夫定理成立所需要的假设。因此，如果需要针对误差项异方差且相关的情况，需要新的估计方法。

模型设定为：

是误差项的协方差矩阵，如果是，便是普通的线性回归模型。在计量经济学中，非对角线型的协方差矩阵多出现在时间序列模型中。我们通过调整回归模型使得满足高斯马尔科夫条件，这种方差成为广义最小二乘法(GLS).

7.2 GLS估计量

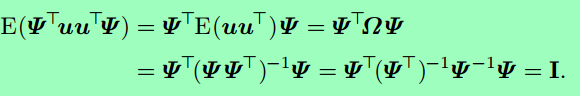
转换通过一个n维方阵，满足以下等式：

将原来的回归模型左乘，得到：

再利用传统方法得到GLS估计量：



推理可知GLS模型的误差项满足独立同方差假设：



同样可求得GLS估计量的方差：



GLS估计量的有效性

GLS估计量也可以被定义为以下矩条件的解（x与残差项正交）：



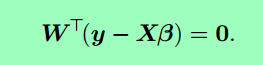
以下矩条件与最小化GLS估计标准[[1]](#footnote-1)的一阶条件相同。由于GLS估计量也是据估计量，可以将其与矩估计量做比较。

注：常用的矩条件 p257

1. 自变量与残差项正交

2. 误差项期望值为0

常见的MM对于线性回归的据估计量被定义为一个n\*k 的矩阵W，满足条件：



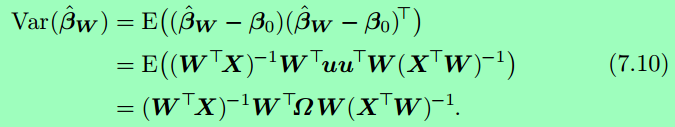
由于有k个方程，k个未知数，我们可以解方程得出MM估计量：



而GLS估计量是MM估计量的一个特例，其中。在特定的条件下，MM估计量是模型的无偏估计量，并且无需进行渐进分析。通过变换可以发现：



因此，beta\_w的协方差矩阵是：



当W=X的时候，这就是一个三明治协方差矩阵。

GLS估计量的有效性可以通过显示矩估计量的协方差矩阵与GLS估计量的协方差矩阵是半正定矩阵（positive semidefinite matrix）来证明。即证明以下矩阵是一个正定矩阵（练习7.2证明）：



GLS估计量比MM估计量更加有效，又因为OLS估计量本身就是矩估计量的一个特例，因此相比OLS估计量，绝大多数情况下GLS估计量将会更加有效，至少不会效率更低。

7.3 计算GLS估计量

是一个很大的矩阵，会带来很大的计算压力，但是如果有如下关系，而n维方阵已知的话可以直接用代入计算过程：



**WEIGHT LEAST SQUARES**

当误差项异方差但是不相关的时候计算相对较容易，因为这时矩阵是对角阵，其逆矩阵也是对角阵，也能够很容易的计算出来，这种特殊的GLS估计被称为WLS，就像是对不同的观察值进行加权，误差项方差大的观察值给予的权重小，反之亦然。有很多种确定观察值权重的方法可以采用。

7.4 可行GLS估计（Feasible Generalized Least Squares）

正常情况下协方差矩阵并不知道，这就让GLS估计量无法计算。然而，在很多情况下可以假设其以一种方式取决于某一个参数，这就使得我们能够估计,这种估计方法被称为可行GLS估计。

7.10 面板数据模型



两种最流行的处理面板数据的模型被称为error-components models. 这种模型的思想是将看成是两到三种独立的冲击，可将之表示为：



其中会影响t时期所有的观察胡，而会影响i类样本的所有观察值，只影响i类样本的所有观察值，每种成分的误差项对应着其下标都各自独立。

为了对误差项成分模型进行估计，和需要被确定为是固定的还是随机的。如果被认定为固定模型，那么它们需要被看成参数进行估计，常使用虚拟变量进行OLS估计。如果被看做随机效应，那么就必须估计出上述模型，然后使用GLS方法进行估计。

固定效应估计(Fixed-Effects Estimation)

第八章 IV估计

误差项与解释变量无关是很强的假设，很多时候并不能成立。但是我们仍然可以找到一个信息集，既满足信息集中的解释变量与误差项无关，也满足：



这个时候需要利用IV估计方法来解决这个问题。

8.2 自变量和误差项之间的相关性

两种常见的误差项和自变量有相关性的情况：

1. errors in variables. 测量误差；

2. simultaneity, 当两个以上内生变量被联立方程所决定；

**Errors in Variables**

当存在测量误差时，得到的是有偏和不一致的估计量。

**Simultaneous Equations**

经济理论表明两个以上的内生变量会被同时决定，在这种情况下，所有的内生变量都可能与误差项相关。意味着这些变量并不会在OLS模型中准确的出现。

书中用供给和需求的例子说明，只要有同时被线性方程决定的估计量，那就必然会导致误差项和内生变量之间的相关性。当存在多重联立方程的时候，如果我们要估计出由许多方程组成的系统，我们需要用到12章中的方法，如果只是需要估计出单一方程的话，我们可以用IV估计方法。

为了解决上述两种可能导致内生性问题的情况，我们需要找到一种新的估计方法来解决内生性问题，即IV估计方法。

8.3 IV估计

简单IV估计

简单IV估计量





the generalized IV estimator

过度识别：IV数量超过了需要估计的参数，当识别不足的时候，矩条件没有唯一的解。

当过度识别时，为选出最优的IV，需要满足以下条件：

1. 矩阵秩为k

2. asyptotically deterministic

3. 矩阵J应最小化所估计系数的协方差矩阵（这里的最小化怎么定义）

满足以上条件的IV矩阵是, 得到的IV估计量为



称之为GIV estimator

两阶段最小二乘法（Two-Stage Least Squares）

1. X对W做线性回归，得到拟合值；

2. 第一阶段的拟合值作为第二阶段的自变量矩阵，对Y进行回归，估计出来的系数即是IV估计量；

第九章 GMM估计（The Generalized Method of Moments）

1. 最小化误差项的方差 [↑](#footnote-ref-1)