

仔细阅读《A SURVEY OF ECONOMETRIC METHODS FOR MIXED-FREQUENCY DATA》
笔记记录在：

NOTE - A SURVEY OF ECONOMETRIC METHODS FOR MIXED-FREQUENCY DATA

1 背景介绍 [Introduction - background](#)：不平衡的计量模型研究越来越多，原因是因为许多宏观经济数据公布频率不相同。

在本文中，研究者关注文献的最新发展，特别是介绍了MIDAS方法及其最近的扩展，经典和贝叶斯框架下的MF-VAR，以及考虑到数据集的混合频率和[粗糙边缘](#)的不同因素模型（多元因素模型）。

2 混频模型

典型的[回归](#)模型是用相同频率抽样的变量来拟合的。为了确保相同的频率，处理时间序列数据的研究人员要么将[高频观测数据聚合到可用的最低频率](#)，要么将[低频数据插入到可用的高频数据中](#)，来进行回归。

但是这样进行处理，就可能会丢失可用的信息，或在数据中插入了错误的信息。因此研究人员提出了直接使用的混合频率数据模型。

- 早期处理混合频率数据的方法之一侧重于预测并依赖于[桥式方程](#) 例如Bafigi, Golinelli, Parigi(2004), 连接低频变量和时间聚合指标的方程。高频指标的预测由具体的高频时间序列模型提供，然后将预测值聚合并代入桥式方程，得到低频变量的预测。
- 在2.2节,2.3节中，我们更详细地概述了一种最新的、最具竞争力的[单变量方法](#) 即由 Ghysels、Santa-Clara和Valkanov(2004)最初提出的混合数据抽样方法。混合数据采样(MIDAS)模型处理在不同频率下采样的序列，其中分布式滞后多项式用于确保简约的规范。早期的MIDAS研究集中在金融应用上，如Ghysels, Santa-Clara和Valkanov(2006)，而最近这种方法已被用于预测宏观经济时间序列，其中通常由月度宏观经济和金融指标预测季度GDP增长，如Clements和Galvao(2008, 2009)。
- 文献中另一种常见的方法是用[模型的状态空间表示](#)处理不同频率的数据，将低频变量认为是缺失高频变量的观测值。
应用卡尔曼滤波光滑器（The Kalman filter and smoother）来估计缺失的观测值（估计高频变量）并产生预测。
- 此外，还有对低频和高频序列的动态进行了联合分析。
目前最引人注目的方法之一是Zadrozny(1988)提出的直接估计VARMA模型的方法，该模型具有不同频率的采样序列。
以同样的方式，Mariano和Murasawa(2010)设置了混合频率VAR (下面简称MF-VAR)，即他们引入了部分潜在时间序列的VAR模型，并将其转换为状态空间形式，详见2.4节。

在状态空间方法中，我们还可以列出所使用的混合频率因子模型

例如，提取一个未观察到的经济状态，并创建一个新的一致性指标或预测和临近预测的 GDP，参见第2.5.1节中的Mariano和Murasawa(2003, 2010)中的小规模应用，以及第2.5.2节和2.5.3节中的大规模模型的Giannone、Reichlin和small(2008)和Banbura和Rnstler(2011)。

Frale等人(2010, 2011)也采用了类似的方法:与其他研究不同的是，*动态因素模型被应用于一组小数据集，其中变量根据经济理论和制度考虑进行分组，而不是应用于整个信息集*。*然后在状态空间框架内将独立的小因素模型链接在一起*。

结论：

- 最后，在第2.5.4节中，我们回顾了文献并建议将混合抽样计量经济学的两个新分支合并：因子模型 (factor model)和MIDAS方法。Marcellino和Schumacher(2010)引入了Factor-MIDAS，这是一种目前利用大量高频指标中的信息预测低频变量的方法。

2.1 聚合和插值

典型的[回归](#)模型是用相同频率抽样的变量来拟合的。为了确保相同的频率，处理时间序列数据的研究人员要么将[高频观测数据聚合到可用的最低频率](#)，要么将[低频数据插入到可用的高频率数据中](#)，来进行回归。

- **将高频聚合至低频：**

标准的聚合方法取决于变量的存量/流量性质。

通常取股票高频变量在一个低频周期内的平均值，以及flow（现金流）的总和，来代表低频周期。

或

假设高频期的信息反映在最新值中，取较高频率变量的最新的值作为股票和现金流变量，代表整个低频期。

- **将低频数据插入到可用的高频数据**

第二种匹配频率的方法是对低频变量进行插值，这种方法很少使用。有几种不同的插值方法，如Lanning (1986)，Marcellino(1998)和Angelini等人(2006)。一种常见的方法是两步过程：首先插入缺失的数据，然后使用新的增强级数估计模型参数，可能要考虑到分解引起的测量误差。从发现了模型的状态空间表示之后，这两个步骤就方便的在卡尔曼滤波器设置中联合运行。

代码块 代表暂时还没有理解的内容