丰富了MIDAS的举例部分和分布式滞后模型的部分，便于理解：

> ## MIDAS模型
>
> 以论文《Mixed Frequency Data Sampling Regression Models：The R Package midasr》中对MIDAS的理解为例，进行讲解：
>
> - 假定$\{y_t, t \in Z\}$为低频可观测的一元时间序列，其滞后算子用$B$表示，即$B_{y_t} = y_{t-1}$ 。
> - MIDAS 回归涉及到对随机过程$\{x_\tau{}^{(i)}, \tau \in Z\}, i = 0, \ldots, k$的线性投影。
>   其中 $x_\tau^{(i)}$ 为高频可观测序列，即在对应的每一个低频时期$t = t_0$ ，我们观测到$x_\tau^{(i)}$的时期， 其中$m_i \in N$为高频解释变量相对于低频解释变量抽样频率的倍差。
> - 当$m_i = 1$时，MIDAS模型退化为同频模型。
> - 高频解释变量$x_\tau^{(i)}$滞后算子用$L$表示，即$Lx_\tau^{(i)} = x_{\tau-1}^{(i)}$。
>
> ---

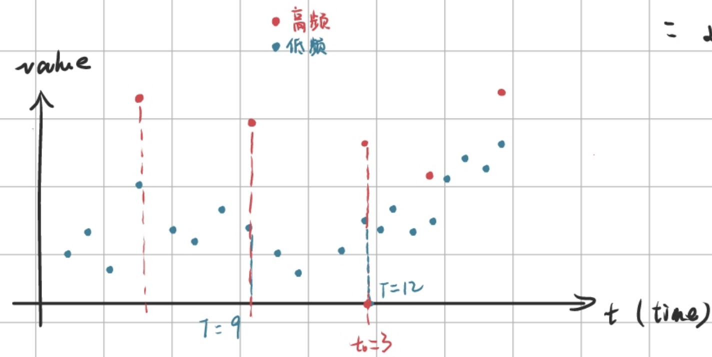
$\{y_t, t\in Z\}$ 低频可观测一元(只讨论一个低频变量)时间序列 $y_t$表示第t个(时间上)低频数据 其滞后算子用B表示 $By_t=y_{t-1}$

$\{x_T^{(i)}, T\in Z\}$ 高频可观测序列 $i=1,2,3,\dots,k$ 表示 $x_T^{(i)}$ 第T个(时间上)高频数据 其滞后算子用L表示 $Lx_T^{(i)}=x_{T-1}^{(i)}$

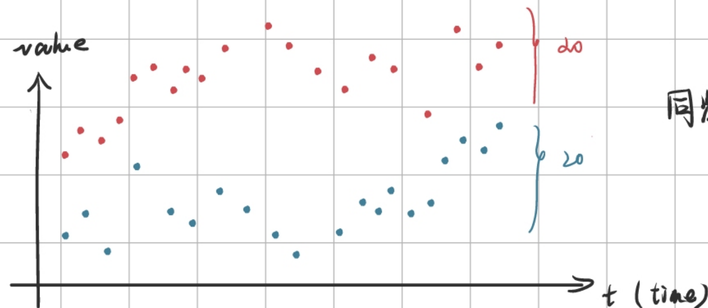
低频时期 $t=t_0$ 时，高频序列 $x_T^{(i)}$ 的时期为 $T=(t_0-1)m_i+1, \dots t_0m_i$，$m_i$为倍差。

比如，高频数据为周数据，低频数据为月数据，那么 $m_i=4$

设 $t_0=3$，低频时期为3时，高频序列 $x_T^{(i)}$ 的时期为 $T=(3-1)\times4+1 \dots 3\times4$
$=2\times4+1=9$ $=12$




$m_i=1$ 时.




同频

---

基于上述设定，MIDAS 模型可表述为如下形式：

$$y_t-\alpha_1 y_{t-1}-\dots-\alpha_p y_{t-p}=\sum_{i=0}^{k}\sum_{j=0}^{l_i}\beta_j^{(i)}x_{tm_{i-j}^{i}}+\varepsilon_t \tag{1}$$

其中要求

$$E(\varepsilon_t|y_{t-1},\dots,y_{t-p},x_{tm_0-l_i}^{(0)},\dots,x_{tm_k-l_i}^{(k)},\dots,x_{tm_k-l_k}^{(k)})=0$$

方程1中的模型可以用通常的时间序列回归方法或使用贝叶斯方法进行估计。但随着模型中滞后阶数的增加，待估参数 $d=p+\sum_i^k l_i$也会快速增加。为了解决这一问题，Ghysels(2004)等建议采用一个充分灵活的函数形式对原参数进行约束，即

$$\beta_j^{(i)}=f_i(\gamma_i,j), j=0,\dots,l_i, \gamma_i=(\gamma_i^{(i)},\dots,\gamma_{q_i}^{i}), q_i\in N$$

通过函数约束可以大大减少待估参数个数，从$d = p + \sum_i^k l_i$减少到$q = \sum_{i=0}^{h_i} q_i$函数化约束虽然可以有效减少待估参数的个数，但也导致模型不再线性化，因此需要采用 NLS （非线性最小二乘法）或极大似然进行估计。

---

为了说明函数化约束的优势，下图分别给出了无约束、正确约束和不正确约束三种条件下均方误差和估计值与真实参数差异的结果。

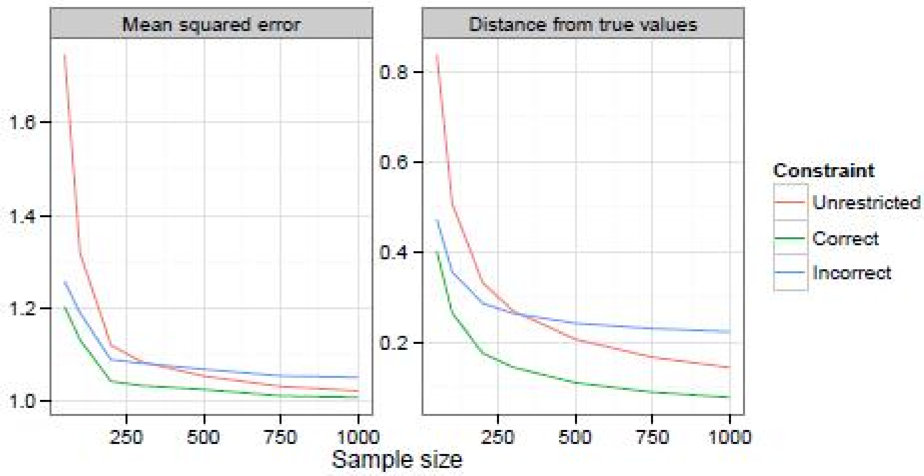



Figure 1: A plot depicting efficiency gains when the correct non-linear constraint is imposed. The left panel plots the average out-of-sample prediction accuracy against sample size. The right panel plots the average Euclidean distance of estimated model parameters to their true values.

> *图1:当施加正确的非线性约束时，描述效率增益的图。左面板绘制了平均样本外预测精度与样本大小的关系。右图绘制了估计模型参数到其真实值的平均欧氏距离。*

由图 1 结果可知，随着样本容量的增加，即使不正确的约束也表现出了较为理想的效果。

---

更为常见的，MIDAS 模型是以一个小巧的形式表述：

$$\alpha(B)y_t = \beta(L)^T x_{t,0} + \varepsilon_t$$

其中$\alpha(z) = 1 - \sum_{j=1}^{p} \alpha_j z^j$，且

$$x_{t,0} := (x^0_{tm_0}, \ldots, x^i_{tm_i}, \ldots, x^l_{tm_l})^T,$$

$$\beta(z) := \sum_{j=0}^{l} \beta_j z^j, \beta_j = (\beta^0_j, \ldots, \beta^i_j, \ldots, \beta^l_j)^T,$$

$$L^j x_{t,0} := x_{t,j} = (L^j x^0_{tm_0}, \ldots, L^j x^i_{tm_i}, \ldots, L^j x^l_{tm_l})^T.$$

- 单阶滞后多项式中，用$l$表示滞后阶数$\beta$的个数$j$的最大值，即$l$为最大滞后阶。

- 如果$\beta(z)$的某些分量的阶较小，则很容易将多项式的某些系数设为零。
  我们要求函数约束对于它的参数的连续二阶导数存在，即$\frac{\partial^2 f_i}{\partial\gamma_i\partial\gamma_i^T}$存在。 函数约束可以随每个变量或频率而变化，因此我们用$\gamma$表示一个受限制的模型的所有参数的向量，用$q = dim(\gamma)$表示参数的总数。

- $\alpha(B)y_t = \beta(L)^T x_{t,0} + \varepsilon_t$中，通过函数约束规范，涵盖了通常的线性(以变量为标准)的MIDAS回归模型的所有变体。

- 而当每个限制函数都是线性映射时，得到一个不受限制的MIDAS回归模型

---

### 频率对齐

首先用矩阵表示法重写模型：
例如，假设每季度观察一次$y_t$，我们想用每月观察一次的变量$x_\tau$来解释它的变化。
因为每个季度有三个月，所以频率倍数$m$在本例中为3。
假设我们假设当前季度和上一季度的月度数据是具有解释力的。这意味着，对于每个季度$t$，我们希望将$y_t$建立模型为季度$t$中观察到的变量$x_{3t}, x_{3t-1}, x_{3t-2}$的线性组合，以及前一个季度$t-1$中观察到的变量$y_{t-1}$和$x_{3(t-1)}$、$x_{3(t-1)-1}$、$x_{3(t-1)-2}$。在矩阵表示法中，本例的MIDAS模型为:

$$\begin{bmatrix} y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} \alpha_1 + \begin{bmatrix} x_6 & \cdots & x_1 \\ \vdots & \vdots & \vdots \\ x_{3n} & \cdots & x_{3n-5} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_5 \end{bmatrix} + \begin{bmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

原来表示的MIDAS模型（1）：

$$y_t - \alpha_1 y_{t-1} - \ldots - \alpha_p y_{t-p} = \sum_{i=0}^{k} \sum_{j=0}^{l_i} \beta_j^{(i)} x_{tm_{i-j}^i} + \varepsilon_t \tag{1}$$

通过将模型写成矩阵表示法，我们可以将高频变量$x_\tau$转换为低频向量$(x_{3t}, \ldots, x_{3t-5})^T$。
我们称这种变换为频率排列。注意，我们要求$x_\tau$的观察次数恰好是$3n$(3的倍数）。

---

另一个例子：假设我们有另一个每周观察的变量，我们想把它添加到模型中。模型(1)中不允许改变频率比，因此我们需要假设每个月恰好有4周。如果一个月不总是有四周，人们可以简单地认为这个模型是固定了一组变量为周滞后(weekly lags)变量。
假设变量$z_\tau$的频率m是12。我们再次使用当前和上一季度的数据来解释$y_t$的变化。这意味着对于$t$季度来说，我们需要将$y_t$用观测到的线性组合的变量$x_{3t}, x_{3t-1}, x_{3t-2}$,和$z_{12_t}, z_{12_t-1}, z_{12_t-11}$来建模，以及季度$t-1$中的$y_{t-1}, x_{3(t-1)}, \ldots, x_{3(t-1)-2}$和$z_{12t}, z_{12(t-1)-1}, \ldots, z_{12(t-1)-11}$，那么模型用矩阵表示为：

$$\begin{bmatrix} y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} \alpha_1 + \begin{bmatrix} x_6 & \cdots & x_1 \\ \vdots & \vdots & \vdots \\ x_{3n} & \cdots & x_{3n-5} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_5 \end{bmatrix} + \begin{bmatrix} Z_{24} & \cdots & z_1 \\ \vdots & \vdots & \vdots \\ z_{12n} & \cdots & z_{12n-23} \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \vdots \\ \gamma_{23} \end{bmatrix} + \begin{bmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

在这个例子中，我们将$x_\tau$对齐为一个向量$(x_{3t}，\ldots，x_{3t-5})^T$,将$z_\tau$对齐为向量$(z_{12t},\ldots,z_{12t-23})^T$。同样的，要求高频变量的观测次数为n的倍数，相乘的因子为对应的频率。

---

现在回到模型(1)的一般情况。我们将高频变量$x_\tau$的频率对齐，通过将其转换为低频向向量$(x^{(i)}_{tm_i},x^{(i)}_{tm_i-1},\ldots,x^{(i)}_{tm_i-l})^T$
那么模型（1）就可以用矩阵表示为：

$$\begin{bmatrix} y_l \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} y_{l-1} & \cdots & y_{l-p} \\ \vdots & \vdots & \vdots \\ y_{n-1} & \cdots & y_{n-p} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix} + \sum_{i=0}^{k} X^{(i)} \begin{bmatrix} \beta_0^{(i)} \\ \vdots \\ \beta_l^{(i)} \end{bmatrix} + \begin{bmatrix} \varepsilon_l \\ \vdots \\ \varepsilon_n \end{bmatrix},$$

其中

$$X^{(i)} := \begin{bmatrix} x^{(i)}_{um_i} & x^{(i)}_{um_i-1} & \cdots & x^{(i)}_{um_i-l} \\ x^{(i)}_{(u+1)m_i} & x^{(i)}_{(u+1)m_i-1} & \cdots & x^{(i)}_{(u+1)m_i-l} \\ \vdots & \vdots & \cdots & \vdots \\ x^{(i)}_{tm_i} & x^{(i)}_{tm_i-1} & \cdots & x^{(i)}_{tm_i-l} \\ \vdots & \vdots & \cdots & \vdots \\ x^{(i)}_{(n-1)m_i} & x^{(i)}_{(n)m_i-1} & \cdots & x^{(i)}_{(n-1)m_i-l} \\ x^{(i)}_{nm_i} & x^{(i)}_{nm_i-1} & \cdots & x^{(i)}_{nm_i-l} \end{bmatrix}$$

$u$是最小的整数,使$um_i - l > 0$ 且 $u > p$即可。
以上将MIDAS回归转换为经典时间序列回归，便于理解，其中所有序列的变量都是相同频率的观察值。