传统利率期限结构模型比较

一、Model

已知利率R和时间t的观测数据，用三个回归模型进行拟合。

- Nelson-Siegel 模型：

$$R = \beta_0 + \beta_1 \frac{1 - e^{-t/\gamma}}{t/\gamma} + \beta_2 \left[\frac{1 - e^{-t/\gamma}}{t/\gamma} - e^{-t/\gamma} \right]$$

其中$\gamma \in [0,10]$。

- Nelson-Siegel Svensson 扩展模型：

$$R = \beta_0 + \beta_1 \frac{1 - e^{-t/\gamma_1}}{t/\gamma_1} + \beta_2 \left[\frac{1 - e^{-t/\gamma_1}}{t/\gamma_1} - e^{-t/\gamma_1} \right] + \beta_3 \left[\frac{1 - e^{-t/\gamma_2}}{t/\gamma_2} - e^{-t/\gamma_2} \right]$$

- Nelson-Siegel & Lasso 回归模型：

$$R = \beta_0 + \beta_1 \frac{1 - e^{-t/\gamma_1}}{t/\gamma_1} + \sum_{i=1}^{10} \beta_{i+1} \left[\frac{1 - e^{-t/\gamma_i}}{t/\gamma_i} - e^{-t/\gamma_1} \right]$$

上式其中 $\gamma_i = i$，且β_1必须为非 0。

二、Data Describe

train： 训练集，共 66 天，对每天数据进行拟合，得到66×3个回归方程。

test： 测试集，共 66 天，用于对上面拟合出的模型进行验证。成绩计算：

$$RMSE = \sqrt{\frac{\sum (y_i - \widehat{y_i})^2}{n}}$$

$$MAPE = \frac{\sum \frac{|y_i - \widehat{y_i}|}{y_i}}{n}$$

自变量为时间t，因变量为利率r。其余字段可以不理。

三、Task

- 拟合出各个模型的未知参数，并把每天的回归参数结果汇总在一个表中。
- 画出每天的拟合效果图。
- 对于模型三，还要画出正则系数α与拟合误差（$RMSE$、$MAPE$）、非零回归系数β个数的关系图。
- 把所有天数的预测函数图汇总在一起。

四、Solving for Nelson-Siegel & Lasso

对于一般回归模型的求解目标函数为误差的平方和最小：

$$\underset{\boldsymbol{\beta}}{\operatorname{argmin}} L(X;\boldsymbol{\beta}) = \frac{1}{2m}\sum_{i=1}^{m}\left(y^{(i)} - \widehat{y^{(i)}}\right)^2$$

其中

$$\widehat{y^{(i)}} = f\left(X^{(i)};\boldsymbol{\beta}\right) = \left[1, X^{(i)}\right]\cdot\boldsymbol{\beta} = \beta_0 + \sum_{k=1}^{d}\beta_k X_k^{(i)}$$

$X^{(i)}$为第 i个样本d维输入向量，$X^{(i)} = \left[X_1^{(i)}, X_2^{(i)}, \dots X_d^{(i)}\right]$；

$\boldsymbol{\beta}$为$d+1$维参数向量，$\boldsymbol{\beta} = [\beta_0, \beta_1, \dots, \beta_d]^T$；

$\widehat{y^{(i)}}$第 i个样本d的拟合函数输出；

$X_k^{(i)}$为第 i个样本第k维的输入值。

Note： $L(X;\boldsymbol{\beta})$中前面的系数不影响参数求解

对于 Lasso 回归，则需要在损失函数中引入$L1$正则。求解目标变成：

$$\underset{\boldsymbol{\beta}}{\operatorname{argmin}} L(X;\boldsymbol{\beta}) = \frac{1}{2m}\sum_{i=1}^{m}\left(y^{(i)} - \widehat{y^{(i)}}\right)^2 + \alpha\|\boldsymbol{\beta}\|_1$$

其中$\|\boldsymbol{\beta}\|_1 = |\beta_0| + |\beta_1| + \cdots + |\beta_d|$

而在本模型中，由于目标函数中β_1必须保留，而后面的自变量作 Lasso 回归，故损失函数为：

$$\underset{\boldsymbol{\beta}}{\operatorname{argmin}} L(X;\boldsymbol{\beta}) = \frac{1}{2m}\sum_{i=1}^{m}\left(y^{(i)} - \widehat{y^{(i)}}\right)^2 + \alpha\sum_{i=2}^{11}|\beta_i|$$

对于模型求解，一般用于梯度下降法，而针对 Lasso 回归，还可以应用最小角回归求解。在此引用第三种巧妙的求解思路：

先对原回归函数中第三项(包括第三项)后面的进行单独的 Lasso 回归：

$$y\_half = \sum_{i=1}^{10}\beta_{i+1}\left[\frac{1-e^{-t/\gamma_i}}{t/\gamma_i} - e^{-t/\gamma_1}\right]$$

回归的因变量还是原来的利率r，得出后面 10 个 Lasso 回归参数

$$\boldsymbol{\beta_{half}} = [\beta_2, \dots \beta_{11}]$$

再把y_half看成新的自变量，再与原模型前两项作一般线性回归：

$$R = \beta_0 + \beta_1\frac{1-e^{-t/\gamma_1}}{t/\gamma_1} + w\cdot y\_half$$

其中w为新回归项y_half的回归系数。

求解出w后再乘以原来的$\boldsymbol{\beta_{half}} = [\beta_2, \dots \beta_{11}]$，这样可以得出所有的回归参数，即：

$$[\beta_2, \dots \beta_{11}] := w\cdot[\beta_2, \dots \beta_{11}]$$

这种求解方法利用了线性回归的线性可加性，使用的前提假设为原模型中各个回归自变量独立不相关，但在本模型中明显不满足这前提假设。但可以用于特征选择。