Combination-forecasting

组合预测

1.预测精度

1.1平均误差 & 平均绝对误差

平均误差：$ME=\frac{1}{N}\sum_{t=1}^{N}e_t=\frac{1}{N}\sum_{t=1}^{N}(x_t-\hat{x}_t)$

平均绝对误差：$MAE=\frac{1}{N}\sum_{t=1}^{N}|e_t|=\frac{1}{N}\sum_{t=1}^{N}|x_t-\hat{x}_t|$

1.2平均相对误差&平均相对误差绝对值

平均相对误差：$MPE=\frac{1}{N}\sum_{t=1}^{N}\epsilon_t=\frac{1}{N}\sum_{t=1}^{N}\frac{x_t-\hat{x}_t}{x_t}$

平均相对误差绝对值：$MRE=\frac{1}{N}\sum_{t=1}^{N}|\epsilon_t|=\frac{1}{N}\sum_{t=1}^{N}|\frac{x_t-\hat{x}_t}{x_t}|$

1.3预测误差的方差和标准差

预测误差的方差：$MSE=\frac{1}{N}\sum_{t=1}^{N}e_t^2=\frac{1}{N}\sum_{t=1}^{N}(x_t-\hat{x}_t)^2$

预测误差的标准差：$RMES=\sqrt{\frac{1}{N}\sum_{t=1}^{N}e_t^2}=\sqrt{\frac{1}{N}\sum_{t=1}^{N}(x_t-\hat{x}_t)^2}$

1.4精度指标

$$\alpha_t=\begin{cases}1-|\frac{x_t-\hat{x}_t}{x_t}| & |\frac{x_t-\hat{x}_t}{x_t}|<1\\ 0 & |\frac{x_t-\hat{x}_t}{x_t}|\geq 1\end{cases}$$

2.组合预测模型

2.1 广义加权平均组合预测模型

对同一预测问题，有n种预测方法。记第t期实际观察值、第i种预测方法第t期预测值和第i种预测方法第t期预测的λ次幂误差分别为$x_t, x_i t, g_{it}$,其中$g_{it} = x_t^{\lambda} - x_{it}^{\lambda}$

（1）λ 次幂误差平方和最小

$$min \quad Q_3 = \sum_{t=1}^{N} g_t^2 = W_n^T G_n W_n$$

$$s.t. \begin{cases} R_n^T W_n = 1 \\ W_n \geq 0 \end{cases}$$

（2） λ 次幂误差绝对值最小

$$min \quad S_3 = \sum_{t=1}^{N} |g_t|$$

$$s.t. \begin{cases} R_n^T W_n = 1 \\ W_n \geq 0 \end{cases}$$

变换成普通的线性规划问题

$$\xi_t = \frac{1}{2}(|g_t| + g_t), \eta_t = \frac{1}{2}(|g_t| - g_t)$$

$$\xi_t + \eta_t = |g_t|, \xi_t - \eta_t = g_t$$

$$min \quad S_3 = \sum_{t=1}^{N} (\xi_t + \eta_t)$$

$$s.t. \begin{cases} \sum_{t=1}^{N} w_i g_{it} = \xi_t - \eta_t, t = 1, 2, \ldots, N \\ w_1 + w_2 + \ldots + w_n = 1 \\ w_1 \geq 0, w_2 \geq 0, \ldots, w_n \geq 0 \\ \xi_t \geq 0, \eta_t \geq 0, t = 1, 2, \ldots, N \end{cases}$$

（3） λ 次幂误差最大偏差最小

$$min \quad M_3 = \max_{1 \leq t \leq N} \{|g_t|\}$$

$$s.t. \begin{cases} R_n^T W_n = 1 \\ W_n \geq 0 \end{cases}$$

变换成普通的线性规划问题

$$\xi_t = \frac{1}{2}(|g_t| + g_t), \eta_t = \frac{1}{2}(|g_t| - g_t)$$

$$\xi_t + \eta_t = |g_t|, \xi_t - \eta_t = g_t$$

$$min \quad M_3 = Z$$

$$s.t. \begin{cases} \xi_t + \eta_t - Z \leq 0 \\ \sum_{t=1}^{N} w_i g_{it} = \xi_t - \eta_t, t = 1, 2, \ldots, N \\ w_1 + w_2 + \ldots + w_n = 1 \\ w_1 \geq 0, w_2 \geq 0, \ldots, w_n \geq 0 \\ \xi_t \geq 0, \eta_t \geq 0, t = 1, 2, \ldots, N \end{cases}$$

Z:新引进的样本期内所有次幂误差绝对值的公共上界

（4） λ 次幂误差极差最小

$$min \quad R_3 = \max_{1 \leq t \leq N} \{g_t\} - \min_{1 \leq t \leq N} \{g_t\}$$

$$s.t. \begin{cases} R_n^T W_n = 1 \\ W_n \geq 0 \end{cases}$$

变换成普通的线性规划问题

$min \quad R_3 = Z_1 - Z_2$

$s.t. \begin{cases} \sum_{t=1}^{N} w_i g_{it} - Z_1 \le 0, t = 1,2,\ldots,N \\ \sum_{t=1}^{N} w_i g_{it} - Z_2 \ge 0, t = 1,2,\ldots,N \\ w_1 + w_2 + \ldots + w_n = 1 \\ w_1 \ge 0, w_2 \ge 0, \ldots, w_n \ge 0 \end{cases}$

Z_1:新引进的样本期内所有次幂误差绝对值的公共上界

Z_2:新引进的样本期内所有次幂误差绝对值的公共下界

2.2广义诱导有序加权平均（GIOWA）组合预测模型

对同一预测问题，有n种预测方法。记第t期实际观察值、第i种预测方法第t期预测值和第i种预测方法第t期预测的λ次幂误差分别为$x_t, x_i t, g_{it}$,其中$g_{it} = x_t^\lambda - x_{it}^\lambda$，第i种预测方法第t期预测精度$\alpha_{it}$。

（1）基于预测λ次幂误差平方和最小的GIOWA组合预测优化模型

$Q_3 = \sum_{t=1}^{N} (x_t^\lambda - sum_{i=1}^{n} w_i x_{\alpha-index(it)^\lambda})^2 = \sum_{t=1}^{N} (\sum_{i=1}^{n} w_i g_{\alpha-index(it)})^2$
$= \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j (\sum_{t=1}^{N} g_{\alpha-index(it)} g_{\alpha-index(jt)} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j G_{ij})$

其中，$G_{ij} = G_{ji} = \sum_{t=1}^{N} g_{\alpha-index(it)} g_{\alpha-index(jt)}, i, j = 1,2,\ldots,N$

$min \quad Q_3 = W^T G W$

$s.t. \begin{cases} R_n^T W = 1 \\ W \ge 0 \end{cases}$

（2）基于预测λ次幂误差绝对值之和最小的GIOWA组合预测优化模型

$min \quad S_3 = \sum_{t=1}^{N} |\sum_{i=1}^{n} w_i g_{\alpha-index(it)}|$

$s.t. \begin{cases} R_n^T W = 1 \\ W \ge 0 \end{cases}$

（3）基于最大λ 次幂误差绝对值最小的GIOWA组合预测优化模型

$min \quad M_3 = \max_{1 \le t \le N} \{|\sum_{i=1}^{n} w_i g_{\alpha-index(it)}|\}$

$s.t. \begin{cases} R_n^T W = 1 \\ W \ge 0 \end{cases}$

（4）基于 λ 次幂误差极差最小的GIOWA组合预测优化模型

$min \quad R_3 = \max_{1 \le t \le N} \{\sum_{i=1}^{n} w_i g_{\alpha-index(it)}\} - \min_{1 \le t \le N} \{\sum_{i=1}^{n} w_i g_{\alpha-index(it)}\}$

$s.t. \begin{cases} R_n^T W = 1 \\ W \ge 0 \end{cases}$

此节模型的求解在第2.1节已经详细介绍

（5）预测期结果

以误差平方和最小为最优准则GIOWA组合预测模型为例。取此模型下样本期内第i种(i=1,2,...,n)单项预测方法的t期（t=1,2,...,N）权重，并做算数平均。在每个预测时点第t期t(t=N+1,N+2,...)对n个权重归一化。得到以误差平方和最小为最优准则GIOWA组合预测模型在预测期的权重向量$(w_1, w_2, \ldots, w_n)^T$。其他模型可以此类推。

使用方法

```
>>> python model.py -h
optional arguments:
  -h, --help            show this help message and exit
  -i INPUT, --input INPUT
                        输入数据文件的位置
  -s S                  是否为诱导模型
  -o OUTPUT, --output OUTPUT
                        输出数据文件的位置
```

基础模型

```
python model.py -i example/example1.xlsx
```

诱导模型

```
python model.py -i example/example2.xlsx  -s yes
```