CTA风格因子计算方法

**步骤一：**通过万德API提出万德商品期货指数日度收盘价：提取的日期为2013-06-28~2018-04-16间的全部交易日，截止18年4月共有48只指数

**步骤二：**整理数据，并导出含5个sheet的xlsx文件：

Sheet1的名称为IDcodes，为48\*2的指数表，第一列为指数的Windcode，第二列为对应中文名称。

Sheet2的名称为DateDay，为1171\*1的日期对应表。

Sheet3的名称为Close，为1171\*48的日度收盘价表，对应行列为DateDay的第一列\*IDcodes第一列。

Sheet4的名称为Volume，为为1171\*48的日度交易量表，对应行列为DateDay的第一列\*IDcodes第一列。

Sheet5的名称为Oi，为为1171\*48的日度持仓量表，对应行列为DateDay的第一列\*IDcodes第一列。

**步骤三：**读取数据并初始化参数。回望周期1为20天(T1)，回望周期2为60天(T2)，滚动周期5天，最小持仓阈值5000手，比例30%。

**步骤四：**记录要剔除的数据。以滚动周期5天为循环间隔，对于每个循环，记录T2个交易日前的交易日无持仓或者近T1个交易日的平均持仓量小于最小持仓阈值5000手的品种，用list记录其位置为delete。

**步骤五：**对于全部48个商品指数，分别计算7个因子：

**因子1：动量因子**

计算方法：

1. 先对于每个指数计算涨跌幅（当前收盘价/T1个交易日前的收盘价），记录为K
2. 剔除delete中记录的对应品种后，从小到大排序（假设剔除完后的list为N，其长度为L），并记录N1为N正数第【L\*30%】个值，记录N2为N倒数第【L\*30%】个值
3. 在K中筛选出小于N1的全部品种K1，大于N2的全部品种K2，再将K1，K2中与delete重复的品种剔除，记为K1’，K2’
4. 将K1’的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的多头部分未来5日的日度收益率，减去，K2’中的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的空头部分未来5日的日度收益率。得到动量因子于未来5日的日度收益率。

**因子2：活跃度因子**

计算方法：

1. 先对于每个指数计算活跃度指标（过去T1个交易日的平均交易量/过去T2个交易日的平均交易量-1），记录为K
2. 剔除delete中记录的对应品种后，从小到大排序（假设剔除完后的list为N，其长度为L），并记录N1为N正数第【L\*30%】个值，记录N2为N倒数第【L\*30%】个值
3. 在K中筛选出小于N1的全部品种K1，大于N2的全部品种K2，再将K1，K2中与delete重复的品种剔除，记为K1’，K2’
4. 将K1’的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的多头部分未来5日的日度收益率，减去，K2’中的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的空头部分未来5日的日度收益率。得到动量因子于未来5日的日度收益率。

**因子3：波动率因子**

计算方法：

1. 先对于每个指数计算波动率指标（过去T2个交易日的日度收益率的标准差），记录为K
2. 剔除delete中记录的对应品种后，从小到大排序（假设剔除完后的list为N，其长度为L），并记录N1为N正数第【L\*30%】个值，记录N2为N倒数第【L\*30%】个值
3. 在K中筛选出小于N1的全部品种K1，大于N2的全部品种K2，再将K1，K2中与delete重复的品种剔除，记为K1’，K2’
4. 将K1’的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的多头部分未来5日的日度收益率，减去，K2’中的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的空头部分未来5日的日度收益率。得到动量因子于未来5日的日度收益率。

**因子4：相关度因子**

计算方法：

1. 先对于每个指数计算相关度指标（某一指数对应的指标为其过去T2个交易日的日度收益率与其余全部指数的相关性之和），记录为K
2. 剔除delete中记录的对应品种后，从小到大排序（假设剔除完后的list为N，其长度为L），并记录N1为N正数第【L\*30%】个值，记录N2为N倒数第【L\*30%】个值
3. 在K中筛选出小于N1的全部品种K1，大于N2的全部品种K2，再将K1，K2中与delete重复的品种剔除，记为K1’，K2’
4. 将K1’的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的多头部分未来5日的日度收益率，减去，K2’中的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的空头部分未来5日的日度收益率。得到动量因子于未来5日的日度收益率。

**因子5：偏度因子**

计算方法：

1. 先对于每个指数计算偏度（过去T2个交易日的日度收益率的偏度），记录为K
2. 剔除delete中记录的对应品种后，从小到大排序（假设剔除完后的list为N，其长度为L），并记录N1为N正数第【L\*30%】个值，记录N2为N倒数第【L\*30%】个值
3. 在K中筛选出小于N1的全部品种K1，大于N2的全部品种K2，再将K1，K2中与delete重复的品种剔除，记为K1’，K2’
4. 将K1’的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的多头部分未来5日的日度收益率，减去，K2’中的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的空头部分未来5日的日度收益率。得到动量因子于未来5日的日度收益率。

**因子6：极端因子**

计算方法：

1. 先对于每个指数计算极端指标（过去T2个交易日的日度收益率的0.95分位值减去过去T2个交易日的日度收益率的0.05分位值），记录为K
2. 剔除delete中记录的对应品种后，从小到大排序（假设剔除完后的list为N，其长度为L），并记录N1为N正数第【L\*30%】个值，记录N2为N倒数第【L\*30%】个值
3. 在K中筛选出小于N1的全部品种K1，大于N2的全部品种K2，再将K1，K2中与delete重复的品种剔除，记为K1’，K2’
4. 将K1’的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的多头部分未来5日的日度收益率，减去，K2’中的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的空头部分未来5日的日度收益率。得到动量因子于未来5日的日度收益率。

**因子7：投机因子**

计算方法：

1. 先对于每个指数计算投机指标（过去T1个交易日平均交易量/当前交易日持仓量），记录为K
2. 剔除delete中记录的对应品种后，从小到大排序（假设剔除完后的list为N，其长度为L），并记录N1为N正数第【L\*30%】个值，记录N2为N倒数第【L\*30%】个值
3. 在K中筛选出小于N1的全部品种K1，大于N2的全部品种K2，再将K1，K2中与delete重复的品种剔除，记为K1’，K2’
4. 将K1’的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的多头部分未来5日的日度收益率，减去，K2’中的指数品种的未来5日的日度收益率按等权重加权求平均和，得到的空头部分未来5日的日度收益率。得到动量因子于未来5日的日度收益率。