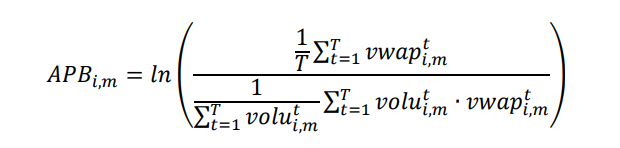
因子回测报告

1. 因子来源与逻辑  
   这次笔试使用的是来自东方证券的-基于量价关系度量股票的买卖压力中的 APB 因子:

我们试图跟踪市场中有已经根据事先的研究判断决定了买卖方向，等待交易时机的投资人。

如果在某一区间内这类投资者持续买入，那么成交量在价格低位相对较大，如果持续卖出，那么成交量在价格高位相对较大。因此，我们就可以根据股票价格和成交量的关系来度量这类投资者的买卖压力，考虑到这类投资者的投资尺度相对较长，买卖持续性较强，买入压力较大的股票后期相对收益可能更高，卖出压力较大的股票后期相对收益可能更低。

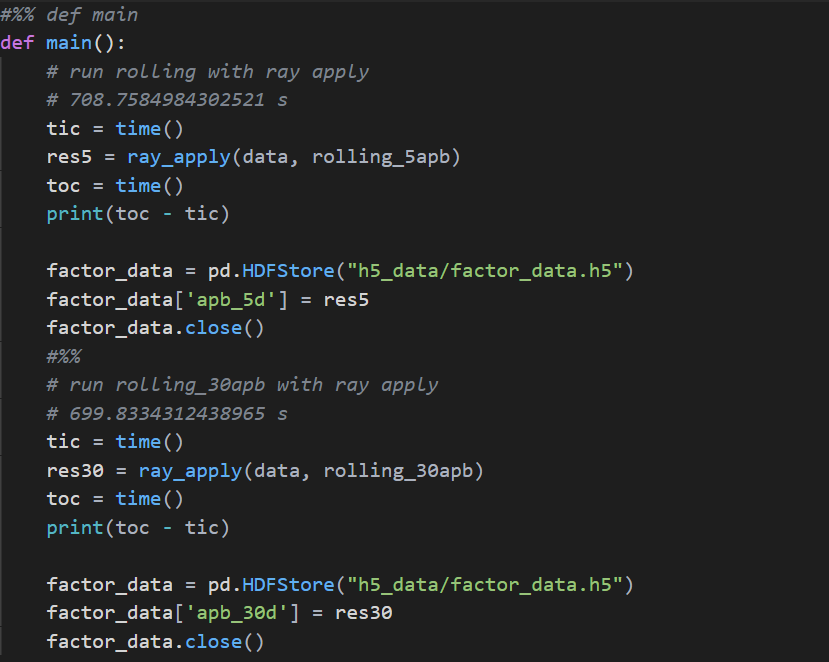
成交量和价格的关系会直接影响到区间内成交量加权价格 vwap 的大小，假设成交量和 价格没有关系，每个交易日成交量一样，那么区间内价格均值就是每个交易日价格的简单平 均，如果价格高位的成交量大，那么区间内 vwap 较高，如果价格低位的成交量大，那么区间内 vwap 较低。因此计算 APB 因子如下:

APB 定义中的分子是各个交易日 vwap 的算术平均值，分母是各个交易日成交量加权的 均值，也就是这个月度区间内的 vwap。如果股票在价格高位成交量大，那么成交量加权的均 值大于算术平均，APB 取值小于零，相反，如果股票在价格低位成交量大，那么成交量加权 的均值小于算术平均，APB 取值大于零，APB 取值越大，股票面临的买压越大，卖压越小。

我们定义回看窗口为一个月的APB因子为: apb\_30d

而我们定义回看窗口为一周的 APB 因子为: apb\_5d

1. 因子计算(factor\_cal.py)

在这次的回测报告中总时长五年的日频数据，我们需要对每一支股票滚动进行计算，以单核计算预估为 1.5 小时左右，因此在这使用多进程计算因子。在 python 的原生 multiprocess 中有一个串行化步骤，因此他在多进程计算大量数据时，加速其实很少。这里使用新的多进程架构 ray(https://docs.ray.io/en/master/)，来加速运算，每种频率的因子计算时间减少为 11 分钟左右

数据与因子的储存都使用 h5 的数据格式，加速读取速度(位于 /h5\_data 文件夹中)

1. 因子报告(factor\_report.py)

因子回测包含 t\_value、ic value、因子收益等信息(位于 /result/factor\_report.html)，其中剔除不能买卖，与 st 股:

* apb\_5d

't均值': 3.4640883541240326

'|t|均值': 17.712733801914954

'|t|>2占比': 0.9240601503759398

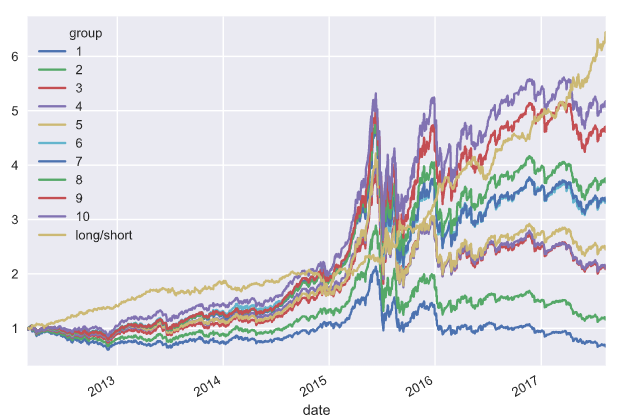
'rank ic均值': 0.03591762431958408

'rank ic标准差': 0.11764236952059204

'rank IC\_IR': 0.3053119761694113

'rank IC大于0的比率': 0.6390977443609023





* apb\_30d

't均值': 3.347242092134932

'|t|均值': 19.415635550473965

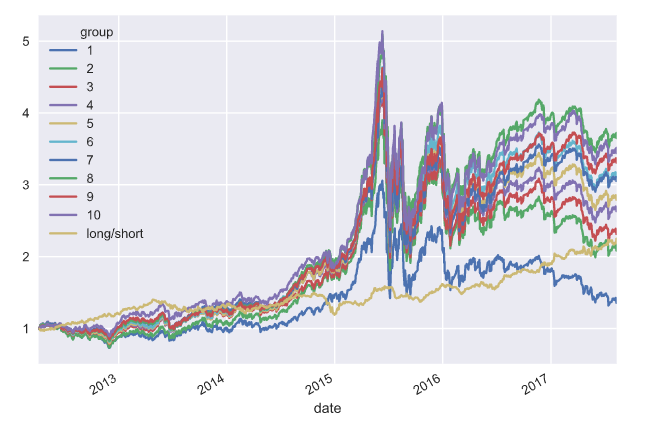
'|t|>2占比': 0.9210727969348659

'rank ic均值': 0.03591762431958408

'rank ic标准差': 0.11764236952059204

'rank IC\_IR': 0.3053119761694113

'rank IC大于0的比率': 0.6390977443609023



可以看到回看一周的因子 apb\_5d 的效果比较好，单调性表现得也比较好

此处的数据皆是使用收盘价的收益作为计算，但实际上的交易我们没办法在因子计算的当天以收盘价买入股票，且为日频换仓，因此年化收益与累计收益仅作参考。更重要的是，该因子的 ic 表现优秀 icir 达 0.3，且 单调性优良，因此我们尝试以apb\_5d 做选股策略的回测。

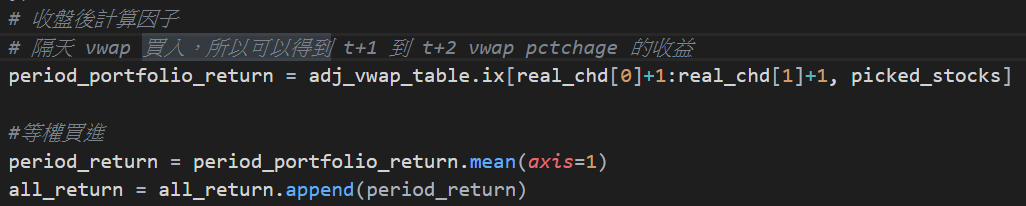
TODO: 此回测没有做 MAD 去极值，以及中性化也没有做，行业中性化也没有。我感觉该因子会跟一些大类因子有很高的相关性。因为这类投资人可能倾向选择优良的股票，以及可能跟一些动量因子有关。

1. 策略回测(trading\_backtest.py)

(报告位于 /result，而每期选出的股票池位于 /stock\_pool)

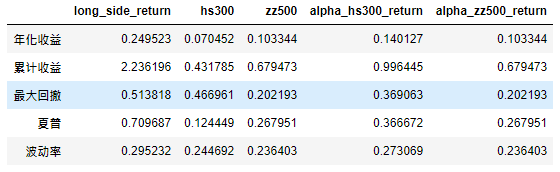
这里我们尝试用 apb\_5d 构建选股策略，我们先尝试月频换仓的策略，考虑每月底的最后一天调仓，如果那天不是交易日则顺延至最近一次的交易日换仓。

此外，考虑我们真实能买到的股票，计算因子当日 (t) 我们要收盘才能计算因子，而隔日我们买进股票 (t+1)，因此我们只能获得 (t+2) 的收益。而我们的收益则是用调整后的vwap 价格计算，以贴近能买到的价格。

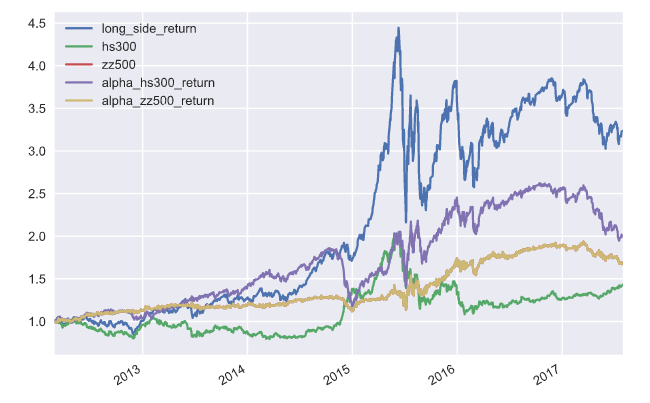


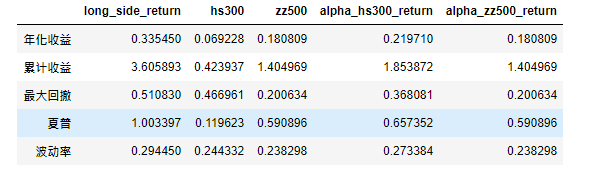
计算完的月频换仓只考虑多头端收益

收益年化为 24%，超沪深300 14%，中证500 10%

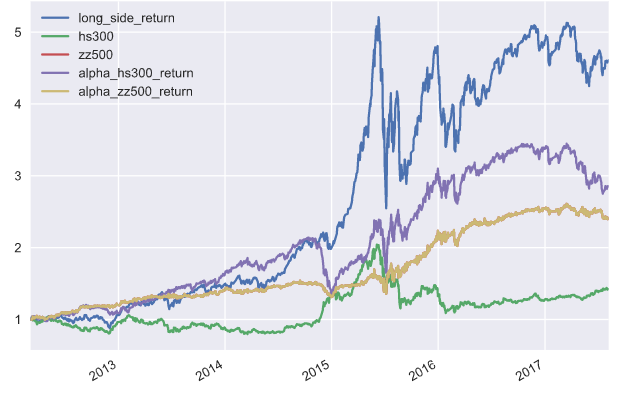






此外也考虑周频换仓的策略: 





有鉴于我们的 abp因子是回看一周的因子，我认为该因子更适合每周换仓的策略，他可以反映近期投资人买进的股票，而周频换仓确实有比较好的效果。

TODO:此次回测因为时间的关系没有考虑到换手率与成本的问题，且持仓以简单的等权做分配。另外还需要考虑到出现一字板的情况。一字涨停或者直接拉涨停的多数情况当日成交量低，利用成交量加权之后反而会拉高apb的值，认为这个股票卖压低，买压高。

还有待加强的地方很多。