

ROE 选股

不稳定问题的原因及改进

一、 不稳定问题的原因

1. 数据库中存在同一时间的重复数据。

如图，在同一日期下，存在相同“发布日期”ANN_DT、但是不同“报告截止日期”的记录。

整批取出数据时，数据排列顺序是不固定的，所以当去重 drop_duplicates 时，每次 drop 的记录不相同，

ANN_DT	REPORT_PERIOD	S_FA_ROE_TTM	STATEMENT_TYPE
20110526	20110331	8.8512	合并报表
20110526	20081231	14.6025	合并报表
20110822	20110630	10.3911	合并报表
20111028	20110930	10.0534	合并报表
20120428	20120331	10.5952	合并报表
20120910	20120630	11.3534	合并报表
20120910	20091231	4.6346	合并报表
20120910	20101231	9.0684	合并报表
20120910	20111231	10.7619	合并报表
20121025	20120930	9.8336	合并报表
20130315	20121231	9.1002	合并报表
20130427	20130331	8.307	合并报表
20130823	20130630	8.0686	合并报表
20131015	20130930	8.5966	合并报表
20140227	20131231	9.6418	合并报表

2. 数据库中缺失数据需要剔除。
3. 计算 ROE_stable 时，没有按时间排序。

二、 改进

1. 提取数据时，选用 Report_Period 为 12 月 31 日的记录（即年报记录）

ANN_DT	REPORT_PERIOD	S_FA_ROE_TTM	STATEMENT_TYPE
20140620	20121231	32.8537	合并报表
20140620	20111231	46.7175	合并报表
20161031	20141231	30.0758	合并报表
20161031	20151231	49.4501	合并报表
20161031	20131231	28.5099	合并报表
20170328	20161231	14.0359	合并报表
20180428	20171231	12.4632	合并报表

发现公布日期仍然很不规律。

采用折中办法，统一将次年 3 月 31 日选为年报公布日期。

2. 剔除缺失数据。

三、 效果

1. 选股数量： 41

2. 选股记录：

600066.SH 600188.SH 603288.SH 600276.SH 002352.SZ 300070.SZ 600309.SH
000540.SZ 601166.SH 002236.SZ 000568.SZ 000157.SZ 000963.SZ 002024.SZ
601877.SH 002415.SZ 000983.SZ 002146.SZ 600570.SH 300144.SZ 002572.SZ
002081.SZ 600177.SH 002739.SZ 000858.SZ 000792.SZ 002602.SZ 600547.SH
002475.SZ 002385.SZ 601933.SH 300059.SZ 601021.SH 000895.SZ 000651.SZ
600406.SH 002411.SZ 300408.SZ 600153.SH 002142.SZ 002558.SZ

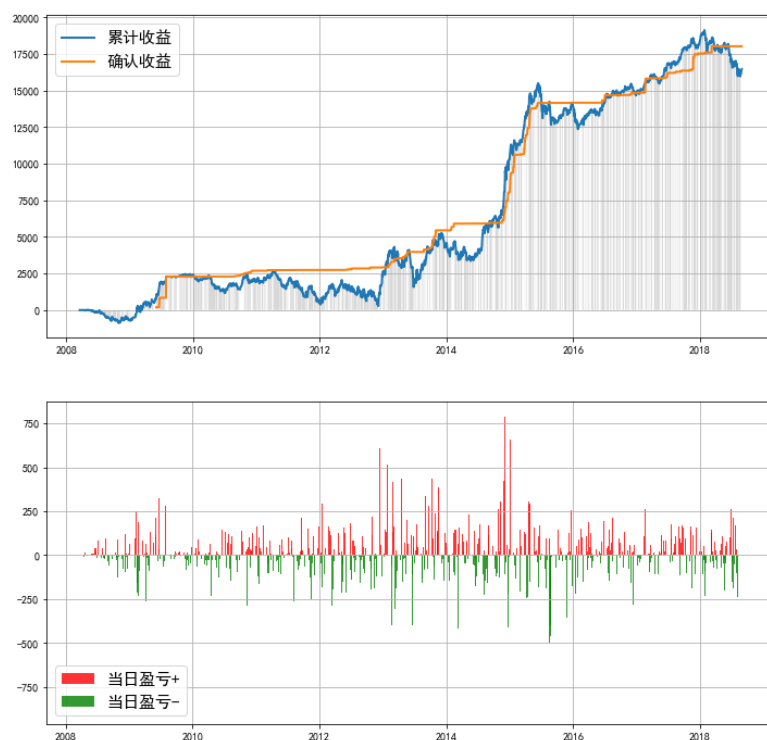
3. 最新选股：

行业	代码	名称
电力行业	0	0
农林牧渔	0	0
钢铁行业	0	0
飞机制造	0	0
生物制药	000963.SZ	华东医药
传媒娱乐	002739.SZ	万达电影
物资外贸	0	0
家具行业	0	0
水泥行业	0	0
农药化肥	0	0
化纤行业	0	0
机械行业	0	0
服装鞋类	0	0
电子信息	600406.SH	国电南瑞
玻璃行业	0	0
酿酒行业	000858.SZ	五粮液
煤炭行业	0	0
交通运输	0	0
金融行业	0	0
船舶制造	0	0
酒店旅游	0	0
环保行业	0	0
商业百货	0	0
汽车制造	600066.SH	宇通客车
食品行业	603288.SH	海天味业
综合行业	0	0
电器行业	0	0
房地产	002146.SZ	荣盛发展
家电行业	000651.SZ	格力电器
次新股	0	0
有色金属	0	0
仪器仪表	002236.SZ	大华股份
石油行业	0	0
其它行业	0	0
化工行业	0	0
医疗器械	0	0
电子器件	0	0
发电设备	0	0
建筑建材	002081.SZ	金螳螂

4. 定投策略回顾：

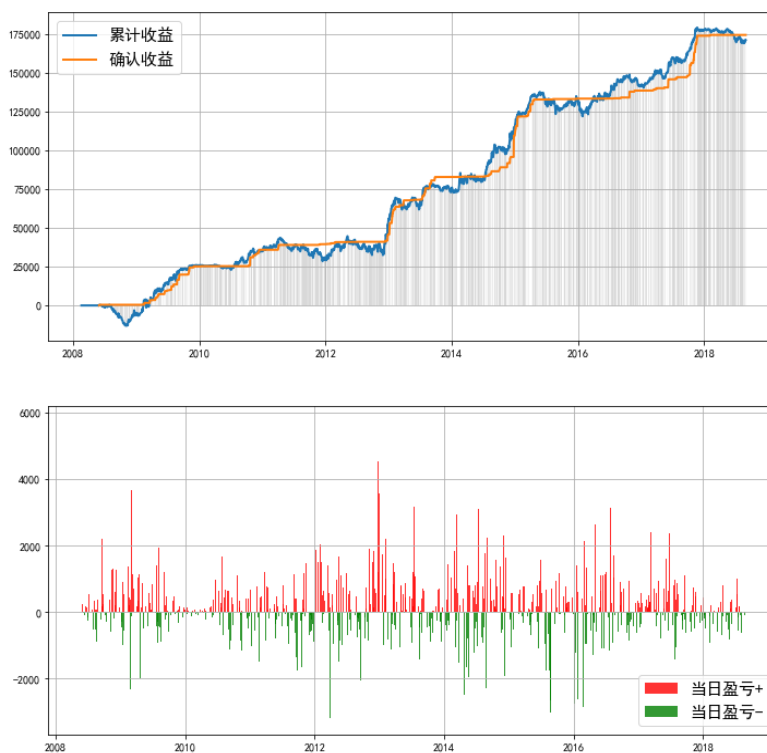
根据在每个行业内选出的股票，择其 roe 最高者进行定投。
每周定投一直持续，直至止盈。止盈后，才重新审视当期该行业选择的股票，
对其进行定投。如果一直没有止盈，定投一直持续。

5. 收益率曲线



四、 比对

之前的收益率曲线



与之前的比对发现，原先的收益率曲线更加平稳。

主要原因是：原先无意中使用了包含了季报的 ROE 数据，因而调整频率较快。而目前改进后，只使用年报数据，在次年 3 月底公布年报数据后决定调仓。相当于

选股频率降到了每年。

其次，原先选出的股票更多，风险更分散。