

E48 土木建筑行业配对交易量化策略简报

1.整体概览

图一为 2019.01.01 到 2020.12.30 的建筑行业指数，2019.01.01 开盘价为 1175.91，2020.12.30 收盘价为 1079.20.行业指数涨幅为-8%。



图 1 880476 建筑行业指数 K 线图 2019.03.01—2019.06.01

2.策略优化

如图 2，002461.SZ 在原始配对交易策略下最主要表现为 β 值低，波动小。但是由于原始策略开平仓条件较为粗糙，导致策略持续性亏损。故策略优化核心思路为**提高策略选股能力，提高收益率。**

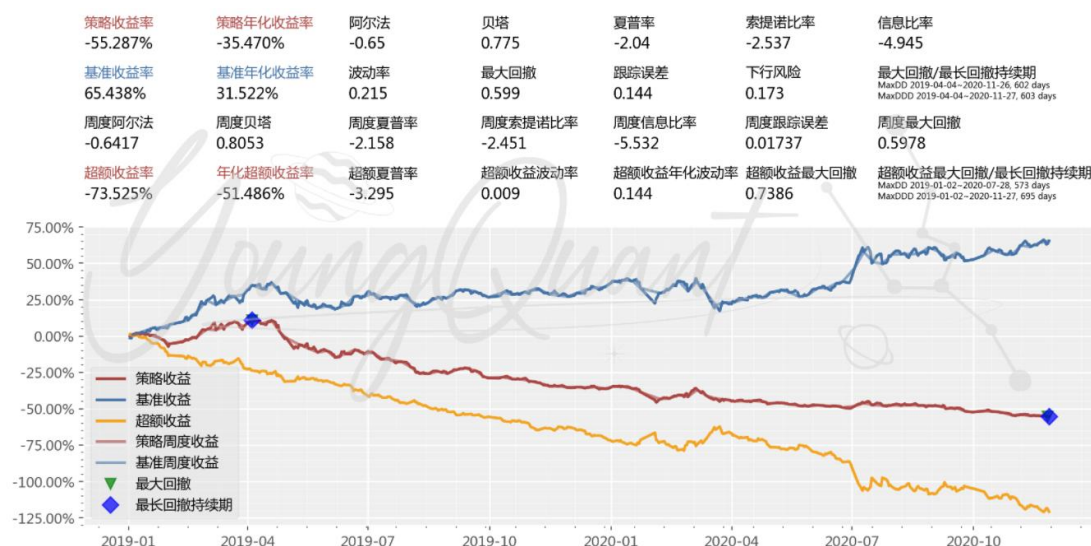


图 2 原始配对交易策略运行结果

步骤一

代码去冗余，删去部分不必要代码。如 `timeDf` 等变量定义了但没有被运用。

步骤二

由于原策略选出近二十个股票对，部分股票协整性并不显著。该原因可能导致亏损。首先尝试按所有股票对的 $pvalue$ 进行排列，选取 $pvalue$ 最小的 5 组进行交易。结果如图 3 所示，发现整体策略收益率有明显上升，从 -55.3% 上升至 -1.4%。



图 3 选取 p 值最小的 5 组股票对

分析: $pvalue$ 越小，股票对协整关系越好。则其均值回复现象出现可能性越大，则投资盈利可能赢越高。

步骤三

延后开仓策略：等价差突破阈值后，反向回归穿越阈值时再建仓。整体上看，由于开仓时价格变化相对不大，收益率变化不大，但是此时股票均值回归趋势更明显，股票盈利可能性越高。

由于 A 股不能够卖空，所以我们只能对股票对被低估的股票进行买入交易。

然而由于日策略回测导致的价格波动较大，本系列回测中没有办法捕捉到穿越阈值的那一瞬间，故修改为**二次判断策略**。即出现连续两次开仓信号才开仓，保证被股票对偏离均值程度更大，提高收益率。

因此在实际操作中，在原策略中对每个股票对增加一个二次判断指标，定义为全局变量: `context.pd_delay = pd.DataFrame`。其内容为布尔数组，初始值为 `len(independent) * len(dependent)` 的 False。

交易信号部分代码如图 4 所示，运行结果见图 5。

```

if price_dep-context.pd_const[stock_de][stock_inde]-context.pd_coe[stock_de][stock_inde]*price_inde>1*context.pd_delay[stock_de][stock_inde] == False:
    context.pd_delay[stock_de][stock_inde] = True
else:
    order_target_percent(stock_de,0)
    order_target_percent(stock_inde,1/order_number)
    context.pd_delay[stock_de][stock_inde] = False
    #突破一倍标准差, 卖出股票y, 买入股票x。
    #context.stock_long[i]=1
    #记录一个多头头寸
    #context.stock_short[i]=0
    #记录一个空头头寸?
#残差序列跌破负一倍标准差开仓
elif price_dep-context.pd_const[stock_de][stock_inde]-context.pd_coe[stock_de][stock_inde]*price_inde<-1*context.pd_delay[stock_de][stock_inde] == False:
    context.pd_delay[stock_de][stock_inde] = True
else:
    order_target_percent(stock_inde,0)
    order_target_percent(stock_de,1/order_number)
    context.pd_delay[stock_de][stock_inde] = False
    #context.stock_long[i]=0
    #context.stock_short[i]=1
else:
    context.pd_delay[stock_de][stock_inde] = False

```

图 4 二次判断策略交易信号



图 5 二次判断策略结果

分析：延后开仓策略效果显著，策略收益率从-1.4%上升至 12.5%， β 值下降 0.25，最大回撤和波动率都有小幅度下降。证明改进有效。

步骤四

在当前策略下，不管何种情况，只要交易信号出现，就会买入 1/5 仓位的股票，显然是不合适的。尤其是仓位中已有某只股票，隔天再次买入，照成不必要损失。所以开仓方式需要针对开仓时的仓位情况进行调整，该股票已有仓位，则增加小部分持仓；没有仓位，则购入 1/5 仓位。

```

if price_dep-context.pd_const[stock_de][stock_inde]-context.pd_cc
if context.pd_delay[stock_de][stock_inde] == False:
context.pd_delay[stock_de][stock_inde] = True
else:
curPosition = get_position(stock_inde).quantity
order_target_percent(stock_de,0)
context.pd_delay[stock_de][stock_inde] = False
if curPosition > 0:
order_target_percent(stock_inde,0.1/order_number)
else:
order_target_percent(stock_inde,1/order_number)

```

图 6 监测仓位代码

运行结果如下（图 7）。



图 7 仓位监测

分析：相较于上一版本，这一版本明显提高了收益。说明了，在检测出股票被低估后，原策略将持续不断地增加持仓。追涨行为将导致收益率的下降。加入仓位监测条件后，策略表现明显上升。

经过参数调整，确定追涨比例为 1/10。



图 8 仓位监测条件优化后运行结果

相较于上一版策略，新策略略微提高了收益率，但是值得一提的是，衡量风险的指标普遍下降，波动率明显下降，抗跌能力明显增强。

分析：相较于上一版本，这一版本**明显降低了波动率**。也说明了，之前的波动率很大程度上是没有加入仓位检测，持续重仓某几只股票导致的。加入仓位监测条件后，策略表现明显上升。从低回撤上看，策略体现出配对交易策略风险中性的特点。

3.总结

将此策略与原始策略比较，收益率提升了约 **77.8%**，风险相关指标明显下降， β 值下降明显。但是夏普比率从 **-2.04** 上升至 **0.84**，综合衡量收益与风险，新策略显著优于原始策略的。尽管在获取超额收益率方面，策略仍然不尽人意，有很多提升空间。相信对开仓条件再增加限制，提高选股能力，在尝试捕捉到更多上涨信号，收益率达到 **50%** 是有可能的。

由于行业自身涨幅限制，回测期内，建筑行业指数收益率 **-8%**，从这点上看，策略对捕捉上涨标的是有明显作用的。

4.附加题

选几个股票池中股票数多的行业，运行策略，结果如下。

从结果看出，策略稳定性较好，在不同行业的表现都比较近似，波动率都维持在低点，回撤控制得比较好。然而收益率在资本市场服务行业有明显上升，可能是资本市场服务业内股票协整关系普遍更强所致。



图 9 B09 有色金属矿采选业



图 10 C15 酒、饮料和精制茶制造业



图 10 J67 资本市场服务