

# Alpha101 研究心得

14331006 毕心然

- **Alpha#11** :  $((\text{rank}(\text{ts\_max}((\text{vwap} - \text{close}), 3)) + \text{rank}(\text{ts\_min}((\text{vwap} - \text{close}), 3))) * \text{rank}(\text{delta}(\text{volume}, 3)))$

其中，**rank(x)**表示某股票 x 值在横截面上的升序排名序号，并将排名归一到[0,1]的闭区间

**delta(x,d)**表示 x 值的最新值减去 x 值在 d 天前的值

**ts\_min(x, d)**, x 中 d 天内最小的值

**ts\_max(x, d)**, x 中 d 天内最大的值

**vwap** 为成交量加权平均价，就是将多笔交易的价格按各自的成交量加权而算出的平均价，若是计算某一证券在某交易日的 VWAP，将当日成交总值除以总成交量即可。

则 **rank(ts\_max((vwap-close), 3))**是找出每只股票三天内成交量加权平均价大于当天收盘价最多的一天的值，然后进行排序和归一化。(1)

**rank(ts\_min((vwap - close), 3))** 是找出每只股票三天内成交量加权平均价大于当天收盘价最少的一天的值，然后进行排序和归一化。(2)

**rank(delta(volume, 3))**为当天成交总量减去 3 天前的成交总量，然后进行排序和归一化。(3)

最终的 alpha 值是 $((1) + (2)) * (3)$ ，其中(1) + (2)的值越大则说明该股票 vwap 与 close 值既不是差距最大也不是最小，是比较居中的位置，因为值过大的话可能说明该股票收盘价低于当天均价，可能处于一个下跌态势，而值过小则可能是因为当天成交量大部分集中在一个较低的价格上，而价格上升是却没什么成交量，属于无量上涨，可能后继乏力，这两者都不希望看到。因此对这两个值取和，这样得到的最大值就是比较稳健的股票。而 (3) 越大则说明该股票近几天的成交量的上升越多，因此整体的 **alpha 值越大说明该股票最近的走势被看好，应进行买入操作。**

**结论 : alpha011 越大越好。**

- **Alpha#81:**  $((\text{rank}(\text{Log}(\text{product}(\text{rank}((\text{rank}(\text{correlation}(\text{vwap}, \text{sum}(\text{adv10}, 49.6054), 8.47743))^4)), 14.9655))) < \text{rank}(\text{correlation}(\text{rank}(\text{vwap}), \text{rank}(\text{volume}), 5.07914))) * -1)$

其中，**rank(x)**表示某股票 x 值在横截面上的升序排名序号，并将排名归一到[0,1]的闭区间

**delta(x,d)**表示 x 值的最新值减去 x 值在 d 天前的值

**ts\_min(x, d)**, x 中 d 天内最小的值

**ts\_max(x, d)**, x 中 d 天内最大的值

**correlation(x, y, d)** x 和 y 两个变量 d 天以来的值的相关系数

**sum(x, d)** d 天以来 x 值的和

**product(x, d)** d 天以来 x 值的乘积

**adv{d}** 为过去 d 天内的平均 dollar volume

以下是我的实现部分：（由于天数只能为整数，我在编码的时候将公式内的值进行了四舍五入）

```
in1 = correlation(vwap, (adv(vwap*volume*100, 10) + 49.6054), 8)
in2 = rank(in1) ** 4
in_left = rank(log(product(rank(in2), 15)))
in3 = correlation(rank(vwap), rank(volume), 5)
in_right = rank(in3)
alpha081 = ((in_left < in_right) * (-1))
```

其中 **in1** 中，**vwap\*volume\*100** 可以算出当日成交总值，即 dollar volume 然后 **adv10** 算出 10 天内的平均成交总值。将这个值加上一个固定值（并不知道怎么来的 49.6054），计算与 **vwap** 在 8 天（由于天数只能为整数，我在编码的时候将公式内的值进行了四舍五入）内的值的相关系数。因此 **in1** 的值越大说明最近 8 天以来，**vwap** 与 10 天内平均成交总值的相关性越大，即与 **volume** 的相关性越小，即 **volume** 可能趋于不变。

**in2** 为将各股票 **in1** 的值进行排序之后再乘一个四次方（也并不知道为什么，可能是扩大差距？）

再将 **in2** 的值进行排序归一化之后，取其 15 天以来的值的乘积，再取对数，再进行一次排序。此时，15 天内 **in1** 值乘积越小的排名越靠后，即 **in\_left** 越小。**in\_left** 值越大则说明该股票交投不活跃，要么流通盘太小，要么大部分流通盘被锁仓。

**in3** 则是计算 5 天内 **vwap** 的排名和 **volume** 的排名的相关性，相关性越大则 **in\_right** 值越大。即 **in3** 越大说明 5 天内该股票的 **vwap** 排名和 **volume** 排名相关性越大，说明交易活跃，带量上涨或者带量下跌。

最终如果 **in\_left < in\_right** 则 **alpha** 值为 -1，若 **in\_left >= in\_right** 则 **alpha** 值为 0。Alpha 为 -1 时说明该股票较为活跃，为 0 时则说明该股票不活跃。

因此该 **alpha** 值为 -1 时比较好。