

2024年9月18日

因子与指数投资揭秘系列二十：构建规则型基本面择时因子

虞堪

投资咨询从业资格号：Z0002804

yukan010359@gtjas.com

高宇飞（联系人）

从业资格号：F03124155

gaoyufei028920@gtjas.com

报告导读：

在量化投资中，规则型的量价因子使用及其广泛。规则型因子是通过一定的数学模型和算法，根据历史数据人为设定的交易规则或策略，从而帮助投资者进行决策。它的优点在于使用相对灵活，可以将投资者的择时想法充分融入因子中。

本文首先介绍了几种常见的规则型因子，如**移动平均线（MA）、双均线策略、通道信号策略**和其他一些**技术指标因子**。我们总结归类了它们构造的原理及一般性步骤。然后，文章探讨了如何仿照构建**规则型的基本面因子**。首先是构造规则型的单因子，它的构造方式较为简单直接，可以直接套用规则型量价因子的构造方式，仅将数据源替换成对应的基本面数据。我们举例分别用布林带和分位数的计算方法构建了库存因子。

接下来，本文分析了如何构造更为复杂的基本面因子，以增强择时能力。一个复杂的规则型基本面因子可以看成由多个单因子，经过一些规则组合而成。它既可以从基本面上下游链条的传导逻辑进行解释，也可以从量化分析中因子组合的角度进行解释。

最后，本文分别以PTA的**库存-利润因子**组合和纯碱的**库存-基差因子**组合为例，展示了规则型基本面因子的实际应用。组合后的规则型因子较单因子有更强的择时能力。我们同时对比了它们的开仓点位分布，组合后的规则型因子对于开仓要求更苛刻，空仓的时间段增加，在出现较大波动的行情时风险控制做的更好。总的来说，构造规则型基本面量化因子可以更好地将基本面择时的逻辑体现出来，但也需要注意避免参数选择较为随意而产生过拟合的风险。

风险提示：注意政策、地缘政治、突发事件等非量化因素的扰动。因子的构建基于历史数据和统计规律，在未来可能会有失效风险。

(正文)

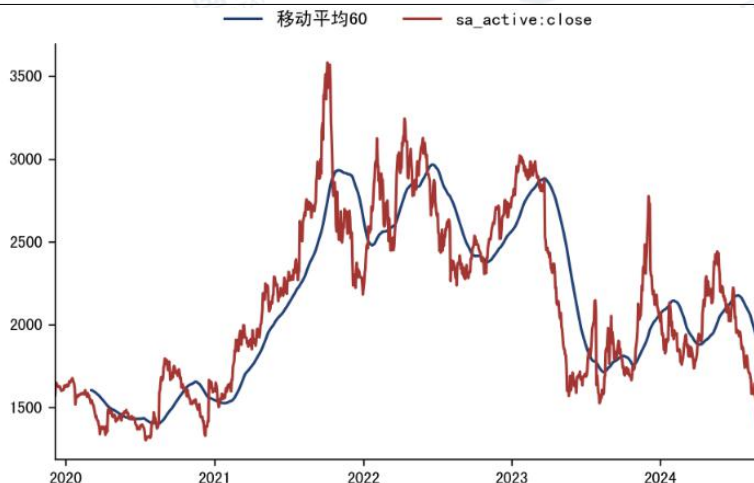
1. CTA 策略中常见规则型因子综述

在量化投资中，不管是权益还是商品市场，规则型因子都是较为常见使用的因子类型。它是通过一定的数学模型和算法，根据历史数据设定的交易规则或策略。这种因子可以对市场的走势进行预测，帮助投资者进行决策。由于择时点或进场的规则是人为设定的，因此使用上较为灵活，也往往可以带来一定的超额收益。我们接下来介绍几种常见的规则型因子：

1.1 移动平均线 (MA)

移动平均线是构造最为基础的因子之一。它的计算方式是求某一个价格序列在过去 N 个周期的平均值，这里 N 个周期可以是过去 N 日，也可以是过去 N 分钟。我们也可以认为它是一种加权平均，过去 N 个值每个权重都是 $1/N$ 。在此基础上，可以赋予过去 N 个值不同的权重，使得近期的数据权重更大，例如**指数移动平均线 (EMA)**。这些不同的移动平均线，反映了价格在过去一段时间的趋势，是趋势策略的重要组成部分。

图 1：纯碱合约收盘价和 60 日移动平均线，移动均线反映了其过去 3 个月的大致走势



资料来源：国泰君安期货研究、同花顺

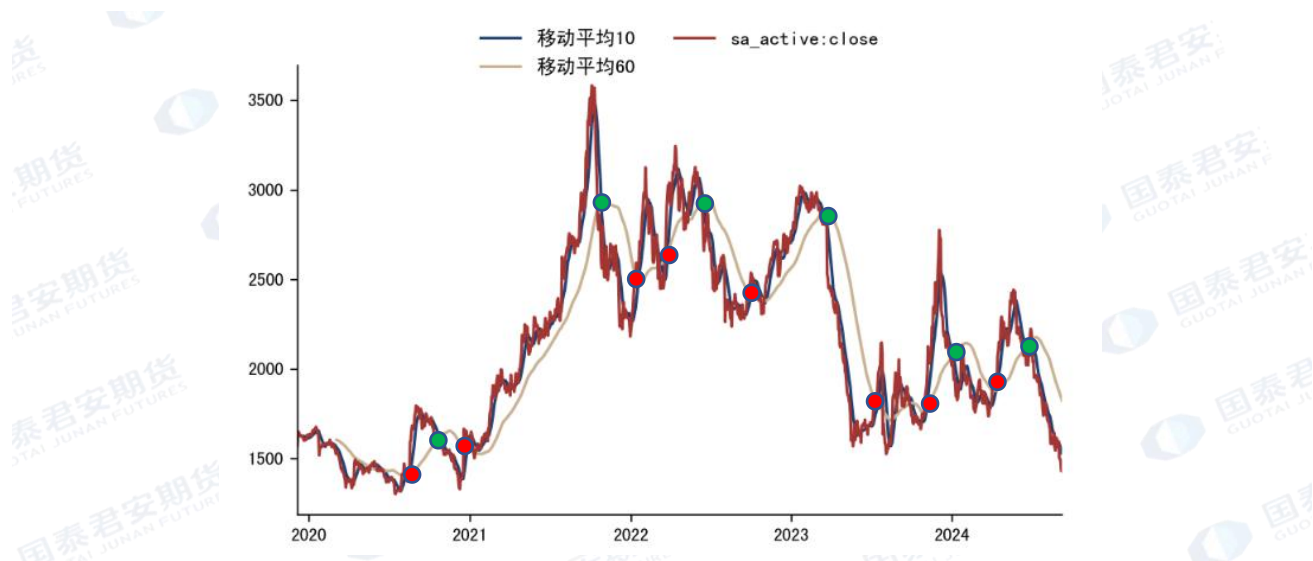
1.2 双均线策略

在移动均线的基础上，我们可以构造简单的**双均线突破**的择时策略。规则设定如下：

1. 分别计算一条短周期均线值，和一条长周期均线值。
2. 当短周期均线向上穿过长周期均线，是择时买入的信号（也称“黄金交叉”）
3. 当短周期均线向下穿过长周期均线，是择时卖出的信号（也称“死亡交叉”）

短周期可以选如 3 日、5 日、10 日等，长周期可选稍长一点的参数，例如 20 日，60 日等，这样我们就构造了一个简单的双均线择时的因子。

图 2：纯碱合约收盘价和 10 日，60 日移动平均线，有多个择时点



资料来源：国泰君安期货研究、同花顺

异同移动均线（MACD）指标可以看作双均线策略的一个简单应用。它的主要构造方式是构造一条短周期的指数移动均线（EMA，周期通常为 12），以及一条长周期的指数移动均线（周期通常为 26）。二者的差值称为离差值（DIF），我们最后再计算 DIF 的指数移动均线（周期通常为 9），称为 DEA 值。其择时原则是：

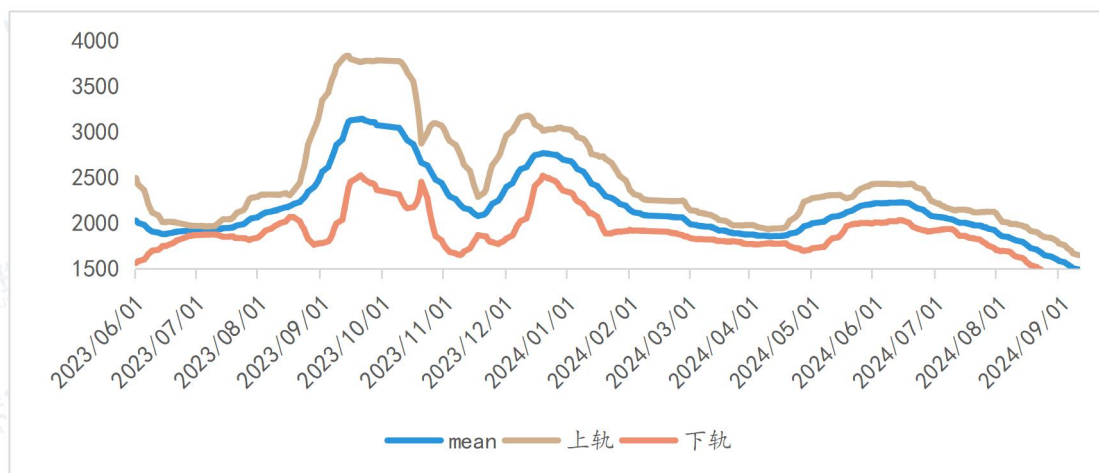
1. DIF、DEA 均为正，DIF 向上突破 DEA，是做多信号。
2. DIF、DEA 均为负，DIF 向下跌破 DEA，是做空信号。

1.3 通道信号策略

常见的通道策略通常由上下两根轨道组成。分别称之为价格的压力位和支撑位。当中间线突破上轨道，则认为是买入择时的信号，而当中间线突破下轨道，则认为是卖出择时的信号。例如常见的**布林带通道**，其构造和择时方式可以分为以下几个步骤：

1. 计算中间线为 N 日价格的移动均线。
2. 计算上下轨分别为中间线加（减）a 倍 N 日移动均价的标准差（a 通常取 2）。
3. 择时信号为，当中间线向上穿过上轨，则代表择时买入（做多）信号；当中间线向下穿过下轨，则代表卖出（做空）信号。

图 3：纯碱 20 日移动均线与布林带通道上中下轨



资料来源：国泰君安期货研究、同花顺

唐奇安通道是另一种常见的通道信号策略，它的计算也相对简单，择时信号直观清晰。具体的构造方式如下：

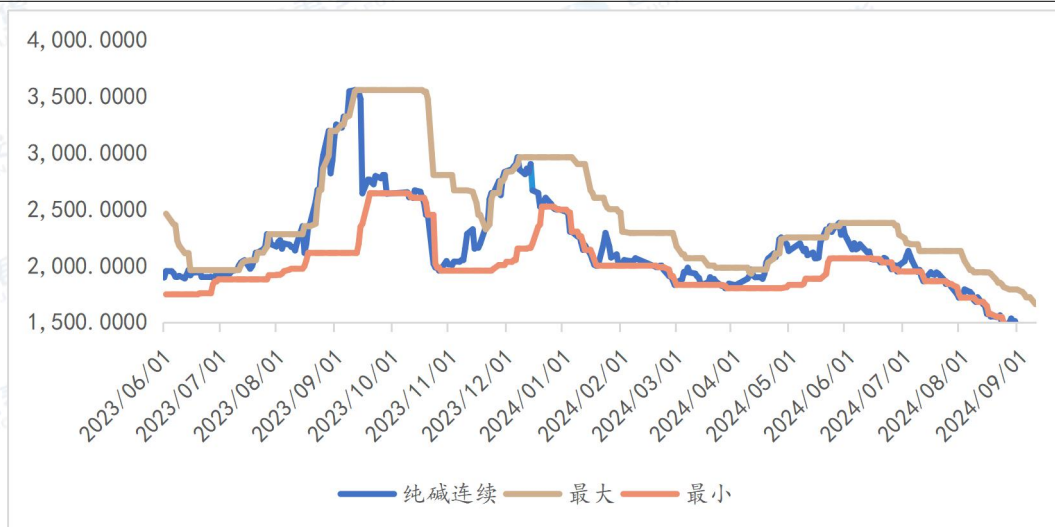
1. 计算上轨道为过去 N 日最高价的最大值，即过去 N 日成交价格的最大值。

2. 计算下轨道为过去 N 日最低价的最小值，即过去 N 日成交价格的最小值。

3. 择时信号为，当价格线向上穿过上轨，则代表择时买入（做多）信号；当价格线向下穿过下轨，则代表卖出（做空）信号。

唐奇安通道在设置止损策略时也非常简单，当在突破上轨做多时，止损点可以设置在下轨处；在突破下轨做空时，止损点可以设置在上轨处。

图 4：纯碱每日收盘价与唐奇安通道



资料来源：国泰君安期货研究、同花顺

1.4 其他技术指标因子

以上介绍的规则型因子，不少也可以归类为技术指标类因子。我们再来看其他一些经常使用的技术指标因子，例如**相对强弱指标（RSI）**。它是通过比较一段时期内的平均收盘涨幅和平均收盘跌幅来分析市场多空头的意向和实力。计算出的 RSI 值，在 0-100 之间，高于 50 为强势市场，反之为弱势市场。当 RSI 值在 80 以上时，代表有超买现象，超过 90 时，代表超买现象严重，未来可能出现反转下跌，此时对应做空操作。相对应地，当 RSI 值在 20 以下时，代表有超卖现象，低于 10 时，代表超卖现象严重，未来可能出现反弹上涨，此时对应做多操作。

同时，长短周期的 RSI 指标也可以像双均线一样，构造择时突破策略。当短周期的 RSI 线向上突破长周期 RSI 线时，是择时做多的信号（黄金交叉）；当短周期 RSI 线向下跌穿长周期 RSI 线时，是择时做空的信号（死亡交叉）。类似的用于分析市场超买（超卖）情况的指标还有**威廉指标（WR）**。

关于技术指标的因子还有很多，我们在《因子与指数投资揭秘系列七：常见技术指标在商品指数上的择时效果》一文已经有较为详细的介绍了。由于并不是本文讨论的重点，故不在此进行一一列举。我们重点讨论在商品期货的基本面量化中，如何借鉴这些规则型因子的构造方式，进行择时操作。

2. 规则型基本面因子的构建

在前文中，我们介绍了一些常见的量价规则型因子的构造方式。对于他们的择时方式，我们可以简单将其总结成以下几点：

1. 若信号满足条件 A，则进行做多（买入）操作；
2. 若信号满足条件 B，则进行做空（卖出）操作；
3. 在其他情况下，继续持有之前的仓位，或空仓操作。

关于条件 A 和 B 的构建方式，主要有 1. 均线或其他价格线的上穿、下跌；2. 构造压力位和支撑位的通道信号策略；3. 计算指标的具体数值，看其所在区间（例如 RSI 指标）等方式。

我们仿照量价规则型因子的构造方式，亦可以构造**规则型的基本面因子**。例如**库存**作为最为重要的基本面数据，一般认为，当商品累库时，价格可能承压下跌；当商品库存连续去化时，价格可能反弹。基于此，我们可以构造**库存的布林带因子**，其择时规则如下：

1. 以库存的 N 日均线作为中轨，N 日均线加减 a 个标准差为上下轨。这里 a 和 N 根据不同的商品品种可以有不同的选择，也可以经过调参后，根据开仓频率选择合适的 a。
2. 当库存值突破上轨，此时认为累库较多，适合开仓做空；反之如果库存值跌破下轨，此时认为去库较多，适合开仓做多。
3. 当库存值处在两轨之间，我们认为库存在合理范围，无明显驱动，空仓或平已有仓位。

对于像**基差**、**仓单**等逻辑较为清晰的基本面因子，也可以采用如上的构造方式。

另一种方式是仿照 RSI 因子的构建方式，将库存数值转化为 0-100 的数值，看其目前所处的位置。例如**库存所处的分位数**，具体择时思路如下：

1. 选择过去 N 日为考察时间窗口，求出当前库存值所处的分位数。

2. 当库存值超过某个值 a (例如 90 分位数), 认为目前累库较多, 适合开仓做空; 反之如果库存值跌破某个分位数 b (例如 10 分位数), 认为目前去库较多, 适合开仓做多。这里参数 a 、 b 和 N 可以根据不同的商品品种有不同的选择, 也可以经过调参后, 根据开仓频率选择合适的值。

3. 当库存分位数处在 a 和 b 之间, 我们认为库存在合理的范围, 无明显驱动, 空仓或平已有仓位。

除了以上的两种构建方式, 还有其他的基本面规则型因子的构造方式, 投资者可以根据自己喜好设置不同的开平仓判断方式和条件, 构造方式较为灵活。

3. 规则型基本面因子的组合, 以两因子为例

3.1 因子构建的逻辑分析

在现实中, 使用单因子进行择时往往效果不佳。以库存为例, 我们建模时通常简单地认为库存对于价格是负向的影响, 但其实它也会受到各种其他因素的影响。如果我们同时考虑利润, 从逻辑上进行分析, 当下游产品利润足够大, 对于原材料的需求较为旺盛, 此时库存是正常去化的一个过程, 因此下游利润大, 库存减少, 价格升高。另一方面, 若下游产品的利润较小, 上游原材料的生产商也可能被动去库存, 此时因下游利润不足, 价格承压。

需要注意的一点是, 上中游产品的加工利润和下游的加工利润逻辑往往相反, 当上中游产品的加工利润较大, 此时供给增加, 价格承压; 反之当上中游产品的加工利润较小时, 此时供给减少, 价格有可能反弹。

我们再来看基差的大小对于库存因子的影响。当基差过大, 此时现货价格远高于期货价格, 边际库存去化, 抛出更多的现货, 有助于基差收敛, 现货承压而期货价格看涨。反过来, 若现货价格远低于期货价格, 此时累库, 流通的现货商品变少, 则现货价格反弹而期货价格承压。

综上所述, 我们不难发现, 可以按照如下的逻辑来构造规则型基本面的两因子:

1. 按照第二节所述的规则型单因子构建的方式, 分别明确因子 a 和因子 b 的做多或做空条件 (阈值), 且在此条件下回测有一定的择时效果。

2. 当 a 和 b 同时满足其开仓做多的条件时, 我们择时做多; 当 a 和 b 同时满足其开仓做空的条件时, 我们择时做空。

3. 在其余条件下均保持空仓或平已有仓位。

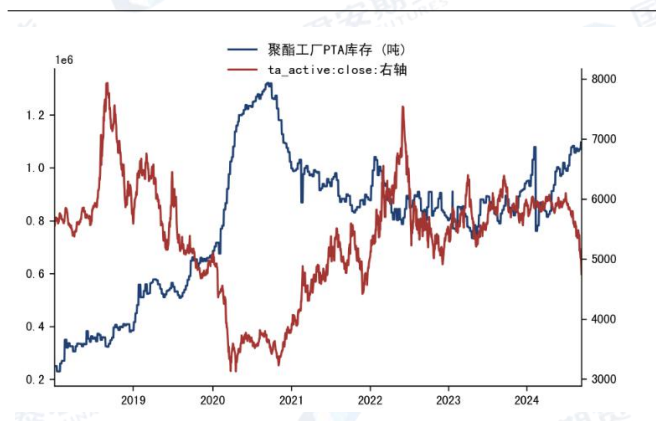
4. 组合后的规则型多因子, 应该比单因子有更强的择时能力。

上面是从基本面逻辑的分析角度出发, 对于构造规则型基本面两因子进行的分析。事实上, 我们也可以从量化的角度进行分析。可以把这个两因子模型视为多因子打分模型的简化版, 我们在之前的系列报告中也介绍过这类模型。在这类模型中, 只有超过一半的单因子给出做多 (或做空) 的信号时, 我们才选择开仓做多 (或做空)。对于只有两个因子的模型来说, 只有它们都给出做多 (做空) 信号时, 我们才会去做多 (或做空), 其他时间处于空仓状态。

3.2 库存-利润因子组合示例, 以 PTA 为例

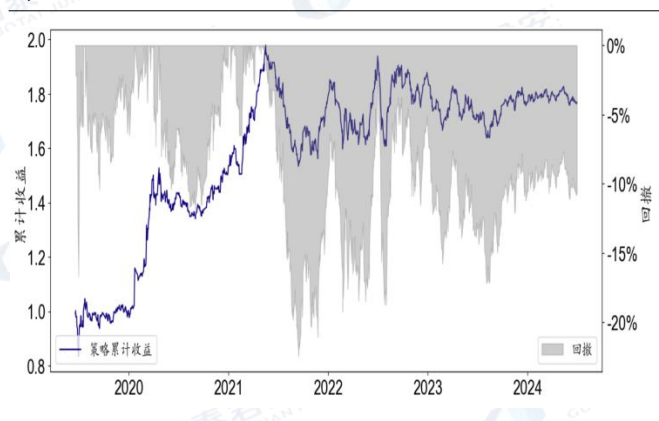
我们以 PTA 的工厂库存和加工利润因子为例, 分别测算其单因子的绩效, 然后再测算按上述组合方法的组合因子的绩效。可以发现, 组合后因子的绩效较单因子均有所提升。

图 5：聚酯工厂 PTA 库存与价格走势



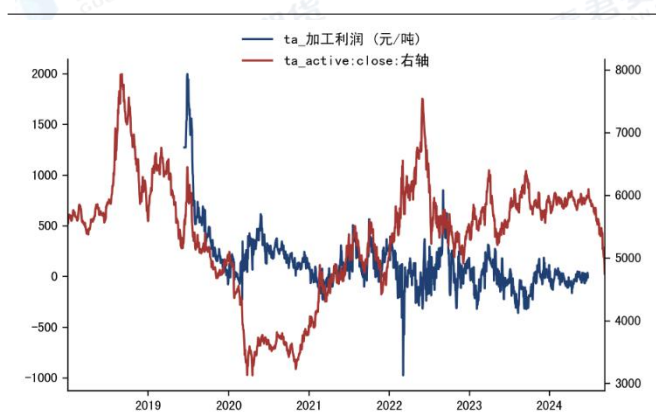
资料来源：国泰君安期货研究、忠朴

图 6：库存单因子回测效果，2019 年至今夏普 0.85，卡玛 0.7



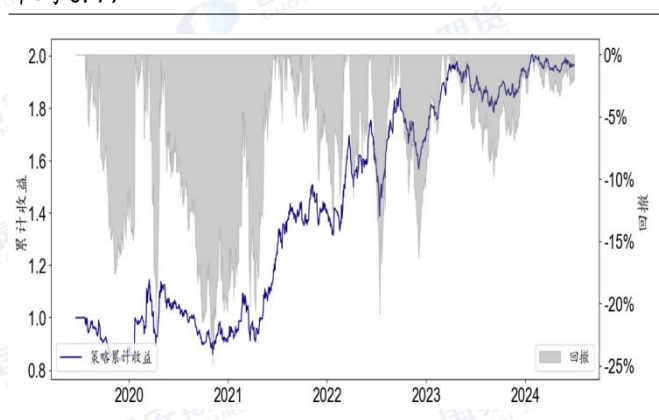
资料来源：国泰君安期货研究、忠朴

图 7：PTA 加工利润与价格走势



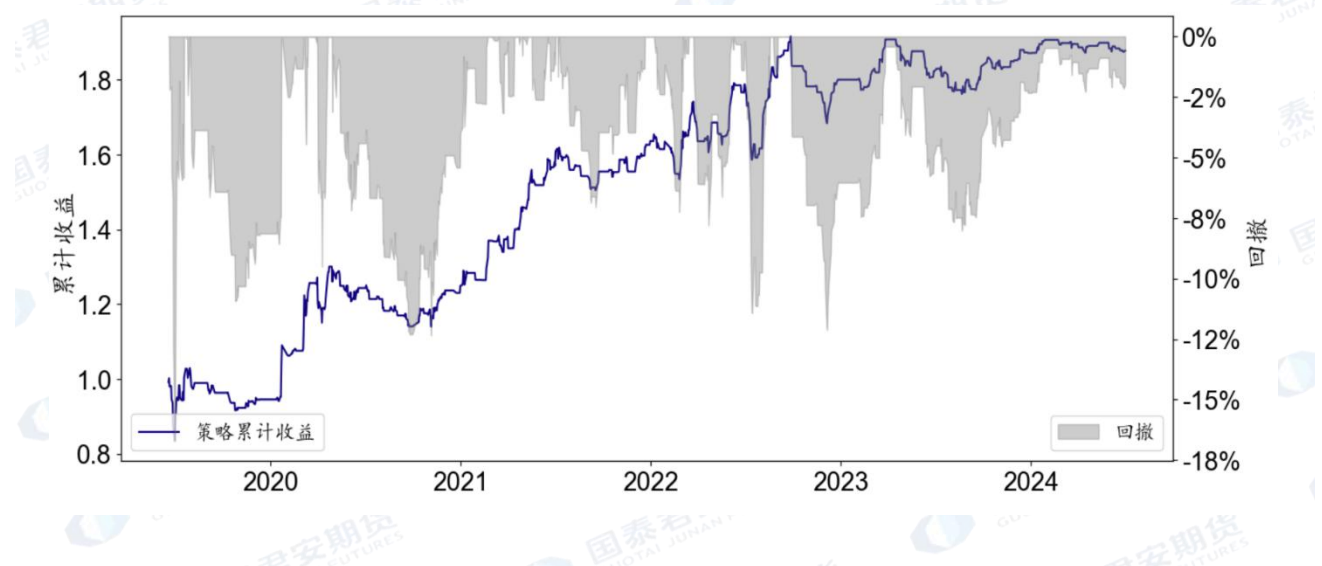
资料来源：国泰君安期货研究

图 8：加工利润单因子回测结果，2019 年至今夏普 0.92，卡玛 0.79



资料来源：国泰君安期货研究

图 9：库存-加工利润合成规则型因子，2019 年至今夏普 1.2，卡玛 1.07



资料来源：国泰君安期货研究

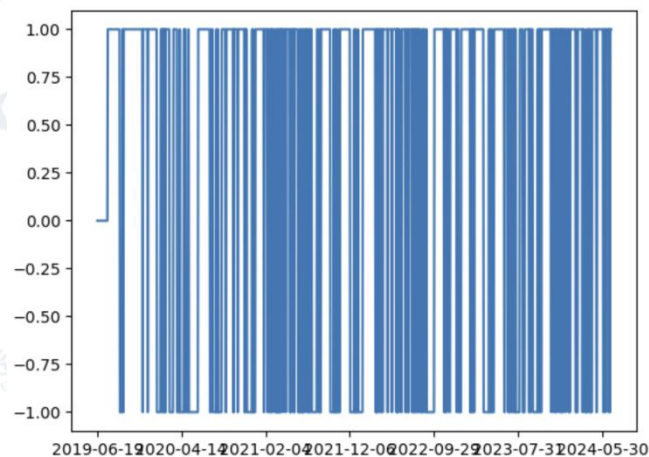
我们再来看择时点分布的情况。单一因子择时开仓频率相对较高，跟踪误差较大。而经过组合后的因子一定程度上降低了开仓频率，控制了回撤风险。下面的择时分布图中，1代表持有多头仓位，-1代表持有空头仓位，0代表目前空仓。可以发现组合后的因子有更多时间处于空仓状态：

图 10：库存因子择时信号分布



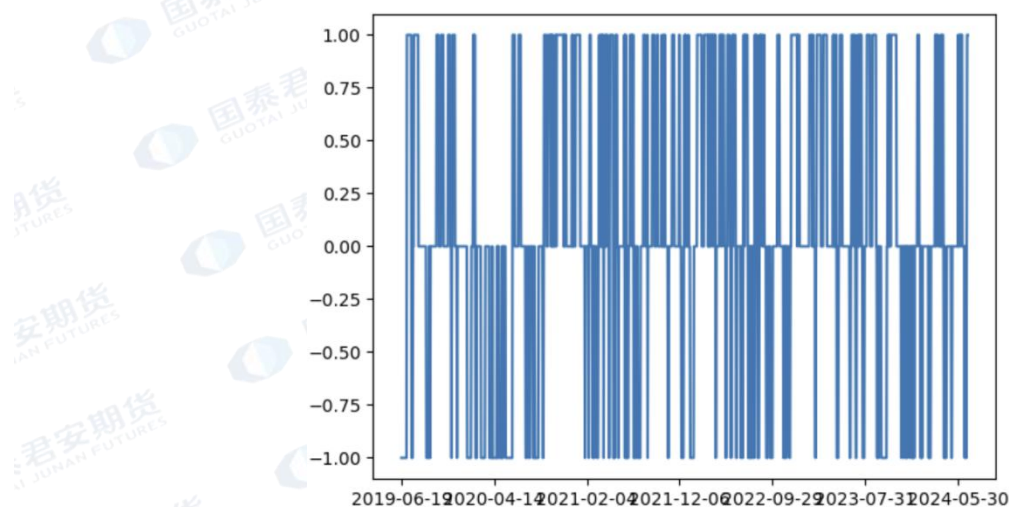
资料来源：国泰君安期货研究、忠朴

图 11：利润因子择时信号分布



资料来源：国泰君安期货研究

图 12：库存-加工利润合成规则型因子择时信号分布

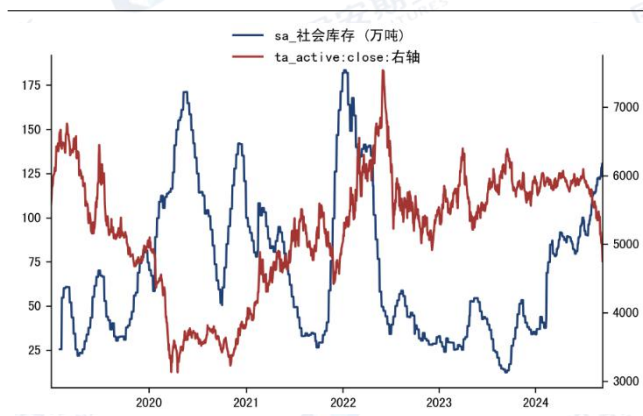


资料来源：国泰君安期货研究

3.3 库存-基差因子组合示例，以纯碱为例

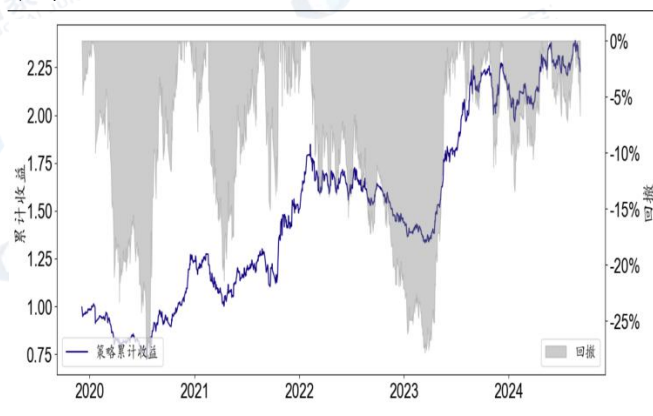
我们再来看另一个例子，以纯碱的库存和基差因子为例，分别测算其单因子的绩效，然后再测算按上述组合方法的组合因子的绩效。基差因子在 2022 年前表现较好，但由于纯碱现货价格近年来波动很大，且一些合约临近交割时基差并未有效收敛，基差因子失效概率较大，结合库存因子可以有效降低风险，控制回撤。可以发现，组合后因子的绩效较单因子均有所提升。

图 13: 社会库存与纯碱价格走势



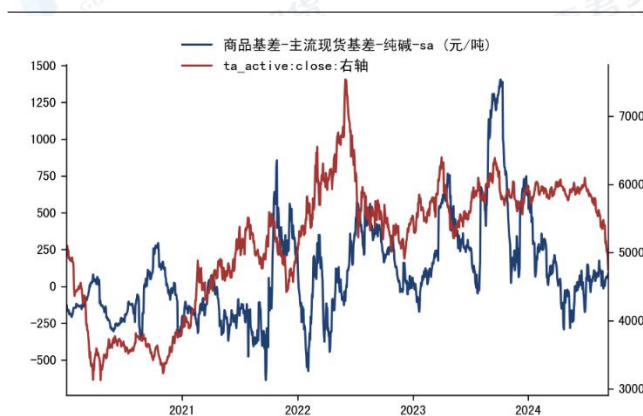
资料来源: 国泰君安期货研究

图 14: 社会库存单因子回测结果, 2019 年至今夏普 1.09, 卡玛 0.94



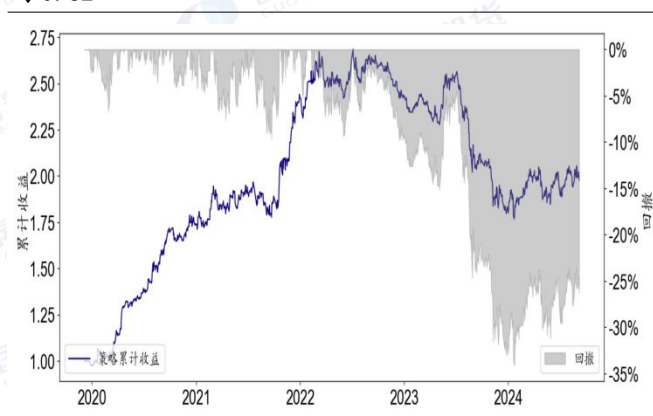
资料来源: 国泰君安期货研究

图 15: 基差因子与纯碱价格走势



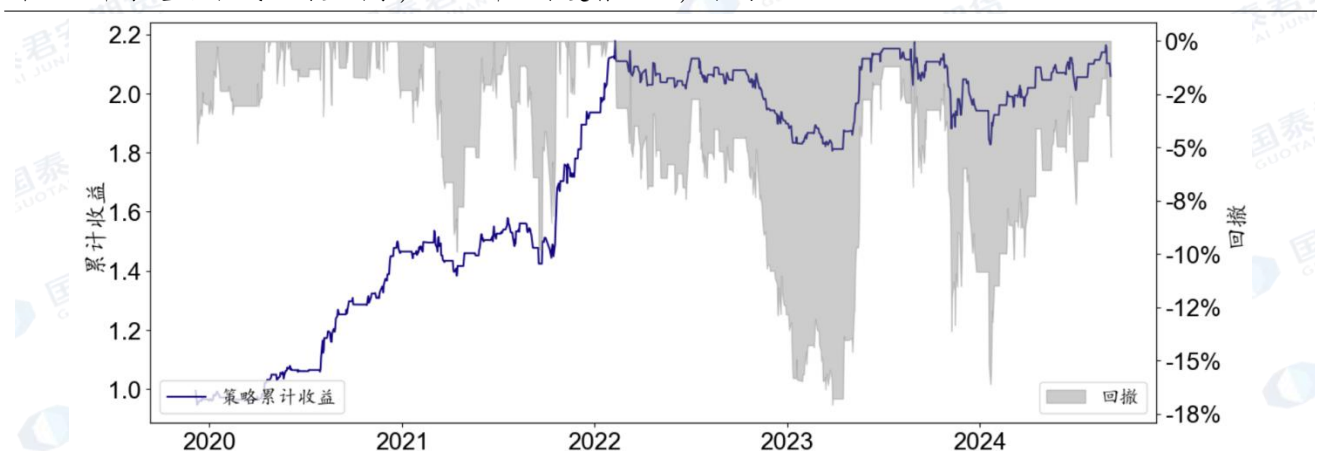
资料来源: 国泰君安期货研究

图 16: 基差单因子回测结果, 2019 年至今夏普 1.18, 卡玛 0.62



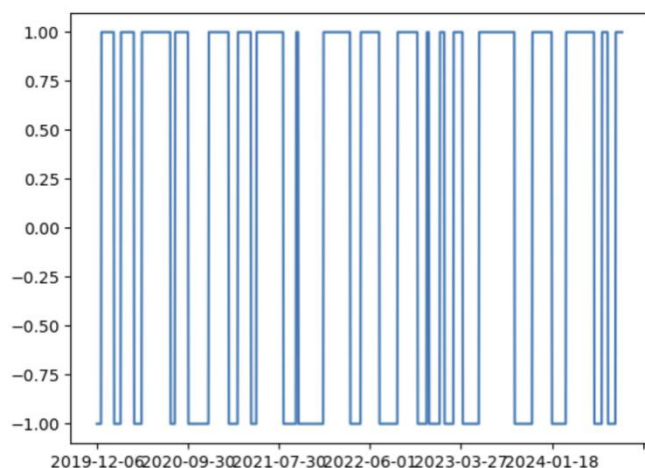
资料来源: 国泰君安期货研究

图 17: 库存-基差合成规则型因子, 2019 年至今夏普 1.53, 卡玛 1.34



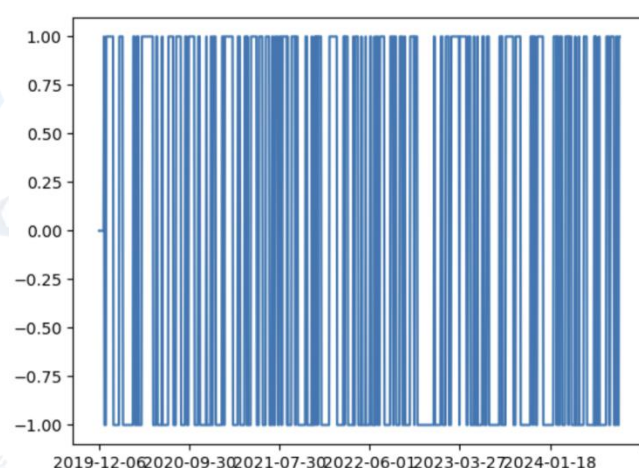
资料来源: 国泰君安期货研究

图 18：库存因子择时信号分布



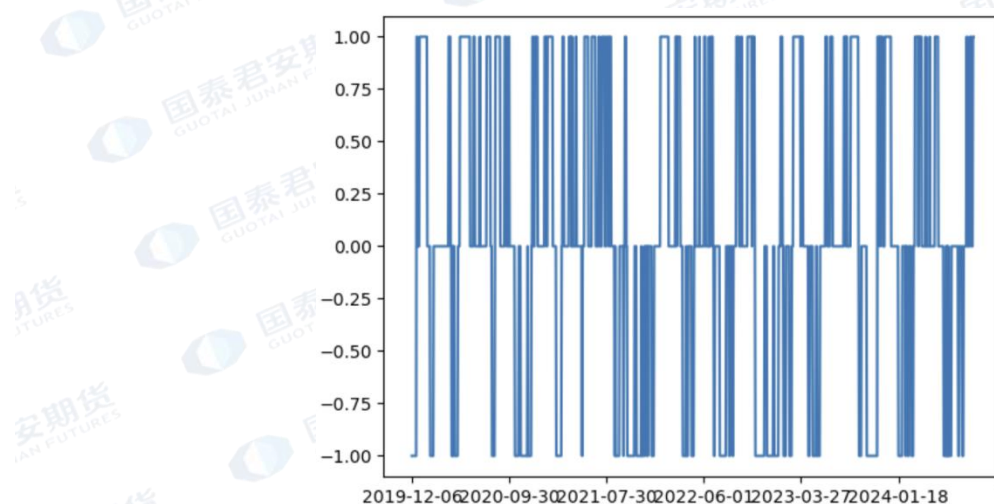
资料来源：国泰君安期货研究

图 19：基差因子择时信号分布



资料来源：国泰君安期货研究

图 20：库存-基差合成规则型因子择时信号分布



资料来源：国泰君安期货研究

4. 总结

我们首先讨论了一些常见的规则型量价因子的构建方式。然后引申到基本面量化上面，讨论如何构建基本面单因子。这一步相对直观清晰，其原理可以仿照规则型量价单因子的构造方式，将原来的价量数据换成对应的基本面数据（如库存、基差、利润等）。我们接着进一步探索相对复杂的基本面规则型因子的构造方式，它可以由多个基本的单因子，结合特定的择时规则构成。本文重点讨论了两因子的情形，首先可以从基本面的逻辑上进行解释，其次可以看成是多因子组合最简单的一种情况。关于三因子及以上的情形，适合直接套用我们之前系列报告中提到的因子组合方法，例如打分法和阶梯强度法，我们在此不再重复。构造规则型基本面量化因子可以更好地将基本面择时的逻辑体现出来，不过也需要注意因为参数选取或条件人为设定过于随机，产生过拟合风险。

国泰君安期货有限公司（以下简称“本公司”）具有中国证监会核准的期货投资咨询业务资格（证监许可[2011]1449号）。

本报告的观点和信息仅供本公司的专业投资者参考，无意针对或打算违反任何地区、国家、城市或其它法律管辖区域内的法律法规。本报告难以设置访问权限，若给您造成不便，敬请谅解。若您并非国泰君安期货客户中的专业投资者，请勿阅读、订阅或接收任何相关信息。本报告不构成具体业务的推介，亦不应被视为任何投资、法律、会计或税务建议，且本公司不会因接收人收到本报告而视其为本公司的当然客户。请您根据自身的风险承受能力自行作出投资决定并自主承担投资风险，不应凭借本内容进行具体操作。

分析师声明

作者具有中国期货业协会授予的期货投资咨询执业资格或相当的专业胜任能力，力求报告内容独立、客观、公正。本报告仅反映作者的不同设想、见解及分析方法。本报告所载的观点并不代表本公司或任何其附属或联营公司的立场，特此声明。

免责声明

本报告的信息来源于已公开的资料，但本公司对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的期货标的的价格可升可跌，过往表现不应作为日后的表现依据。在不同时期，或因使用不同假设和标准，采用不同观点和分析方法，本公司可发出与本报告所载资料、意见及推测不一致的报告，对此本公司可不发出特别通知。本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本报告中所指的研究服务可能不适合个别客户，不构成客户私人咨询建议，客户应考虑本报告中的任何意见或建议是否符合其特定状况。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何直接或间接损失或与此有关的其他损失负任何责任。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。市场有风险，投资需谨慎。投资者不应将本报告作为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向专业人士咨询并谨慎决策。

版权声明

本报告版权仅为本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“国泰君安期货研究”，提示使用本报告的风险，且不得对本报告进行任何有悖原意的引用、删节和修改。若本公司以外的其他个人或机构（以下简称“该个人或机构”）发送本报告，则由该个人或机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该个人或机构以要求获悉更详细信息或进而交易本报告中提及的期货品种。本报告不构成本公司向该个人或机构之客户提供的投资建议，本公司、本公司员工或者关联机构亦不为该个人或机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

除非另有说明，本报告中使用的所有商标、服务标记及标记均为国君期货所有或经合法授权被许可使用的商标、服务标记及标记，未经国君期货或商标所有权人的书面许可，任何单位或个人不得使用该商标、服务标记及标记。