

量化投资在商品期货市场中的策略优化研究

□ 张 达

摘要：本文通过收集和处理 RESSET 与 CSMAR 数据库的沪深 300 期货及指数数据，建立了线性回归模型，并进行了模型验证和结果分析。研究发现，沪深 300 收盘指数对收盘价有显著影响。基于此，进一步研究了量化投资在商品期货市场中的策略优化问题，提出了包括多因子选股、风格轮动选股、趋势跟踪、均值回归、统计套利和股指期货套利等优化方法。提高了投资决策的科学性和准确性。本文的研究为量化投资者在商品期货市场中制定有效策略提供了参考。

关键词：量化投资；期货市场；策略优化

引言

随着金融市场的不断发展，商品期货市场日益活跃。量化投资作为一种基于数学模型和计算机技术的投资方法，在商品期货市场中具有重要的应用价值。通过量化投资，可以挖掘市场中的潜在规律，提高投资决策的科学性和准确性。然而，商品期货市场受到多种因素的影响，如宏观经济形势、供需关系、政策法规等，使得量化投资策略在该市场中的应用面临诸多挑战。在西方投资界，量化投资被归为具有传统投资哲学的投资类型。自其出现起，至今已经历了四十多年的漫长岁月。在这期间，通过大量的实践和数据表明，量化投资是一种能够长期稳定地获取超出平均收益的投资手段。因此，对量化投资在商品期货市场中的策略进行优化研究具有重要的现实意义。

一、研究方法

(一) 数据收集

1. 数据来源。数据来源为 RESSET 金融研究数据库与国泰安 CSMAR 数据库，样本区间均为 2013 年 10 月到 2024 年 10 月。

2. 数据选取。数据选取沪深 300 期货行情数据和沪深 300 指数行情数。对于时间跨度，选取了近 1 年的数据，以保证数据具有足够的样本量来反映市场的长期规律。

(二) 数据预处理

1. 数据检查。确认数据集中的数据类型，确保参与归一化的数据应为数值型数据。如果存在非数值型数据，需要进行转换或处理。检查数据类型的一致性，防止同一数据列中出现多种数据类型混杂的情况，避免在归一化计算时出现错误，查看数据集中是否存在缺失值，缺失值可能会影响归一化的结果，导致数据分布失真。确定缺失值出现的位置和频率，以便采取合适的处理措施，如填充、删除等^[1]。

2. 数据归一化。为了使不同量级的数据能够在同一尺度下进行分析，对数据进行归一化处理。采用最小 - 最大归一化方法，公式为：

$$X_{\text{new}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad \text{公式 1}$$

其中 X 为原数据， X_{\min} 为数据的最小值， X_{\max} 为数据的最大值。

(三) 模型建立

1. 变量设定。以期货价格为因变量，选取可能影响期货价格的因素作为自变量。自变量包括宏观经济变量、行业相关变量等。将这些自变量代入线性回归模型，构建均值回归模型。

2. 模型构建。线性回归模型的表达式为：

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \quad \text{公式 2}$$

其中 Y 表示期货价格， $X_i (i=1, 2, \dots, n)$ 为自变量， $\beta_i (i=1, 2, \dots, n)$ 为回归系数， ε 为随机误差项。

3. 模型验证。进行 F 检验以判断整个回归方程是否显著。F 统计量的计算公式为：

$$F = \frac{(SSR/p)}{(SSE/(n-p-1))} \quad \text{公式 3}$$

其中 SSR 为回归平方和， SSE 为残差平方和。若 F 统计量的值大于临界值，则认为回归方程显著。

(四) 结果分析

对沪深 300 的收盘价与收盘指数经过 SPSS 输出结果进行解释。

如表 1 显示，所有被解释变量的平均离差平方和都为 6916559.679，其回归平方和以及均方都是 6857175.080，剩余平方和都是 59384.598，其均方都是 248.471，在 F 检验统计中的观测平均值为 27597.473，对应的概率 p 值小于 0.05，这表明被解释变量与所有解释变量存在显著的线性关系，所以说

表 1 ANOVA 分析表

| 模型 | | 平方和 | 自由度 | 均方 | F | 显著性 |
|----|----|-------------|-----|-------------|-----------|-------------------|
| 1 | 回归 | 6857175.080 | 1 | 6857175.080 | 27597.473 | .000 ^a |
| | 残差 | 59384.598 | 239 | 248.471 | | |
| | 总计 | 6916559.679 | 240 | | | |

a. 因变量：收盘价

b. 预测变量：(常量)，收盘指数

表 2 回归系数分析表

| 模型 | B | 未标准化系数 | | 标准化系数 | t | 显著性 |
|----|------|---------|--------|-------|---------|------|
| | | 标准错误 | Beta | | | |
| 1 | (常量) | -85.833 | 21.444 | | -4.003 | .000 |
| | 收盘指数 | 1.024 | .006 | .996 | 166.125 | .000 |

a. 因变量：收盘价

明能够建立线性模型。

表 2 中各列数据项的含义依次是偏回归系数、其标准误差、标准化偏回归系数、回归系数显著性检验中统计量的观测值，以及对应的概率、解释变量的容忍度和膨胀因子。基于此，能够开展回归系数显著性检验、写出回归方程。可以发现，在显著性水平为 0.05 的情况下，t 检验的概率 p 值小于 0.05，由此可知解释变量间的线性关系显著。

同时可得回归方程为：

$y = -85833 + 1.024x$

公式 4

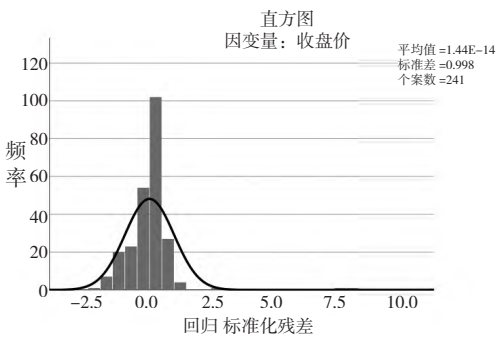


图 1 残差的直方图

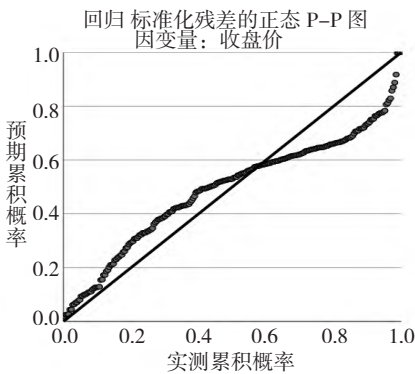


图 2 回归标准化残差的正态 P-P 图

对模型的残差进行分析，检查残差是否符合正态分布假设。若残差不服从正态分布，可能表明模型存在问题，如遗漏了重要变量或模型形式不正确。可通过绘制残差的直方图呈现出钟形曲线的形状，即中间高、两边低且大致对称，这表明残差近似服从正态分布，同时从标准化残差的 P-P 图来看，数据点在图中大致均匀落在一条呈 45 度角度的直线两侧，说明残差服从正态分布。

二、研究结论及原因分析

(一) 研究结论

通过对沪深 300 的收盘价与收盘指数线性模型进行分析，根据分析结果可以得到沪深 300 的收盘指数对沪深 300 收盘价有显著影响。沪深 300 收盘指数可以在一定程度上反映市场对沪深 300 成分股整体的看法和预期，这种整体的市场看法和预期可能会反过来影响投资者对沪深 300 成分股的交易行为，从而间接影响到成分股的收盘价。

(二) 原因分析

1. 投资策略影响。部分投资策略以沪深 300 指数为基准，诸如指数基金投资及若干量化投资策略等。当沪深 300 收盘指数触及某些特定水平或发生特定变动时，此类投资策略或将触发相应的交易举措，从而对沪深 300 成分股的收盘价产生影响。例如，沪深 300 指数若突破某一关键的技术节点，部分量化基金或会依据预设策略，对成分股进行买入或卖出操作^[2]。

2. 板块联动影响。沪深 300 成分股涵盖了不同行业的优质公司，这些公司在一定程度上具有板块联动性。当沪深 300 收盘指数因为某些板块的表现而上涨或下跌时，同板块的成分股可能会受到影响，从而影响其收盘价。例如，如果金融板块的股票在当天对沪深 300 指数的上涨贡献较大，那么其他金融板块的沪深 300 成分股可能会因为板块的整体上涨预期而受到

投资者的关注和买入,推动其收盘价上涨。

三、量化投资策略优化

(一) 选股策略

1. 多因子选股。量化投资中多因子选股策略会参考沪深 300 收盘指数。一些量化模型可能将成分股在沪深 300 指数中的权重作为一个因子。如果一只股票在沪深 300 指数中的权重较高,且该指数近期表现良好,那么这只股票在量化模型中的得分可能会较高,从而更有可能被选入投资组合^[3]。这是因为沪深 300 指数反映了市场对这只成分股的整体评价和预期,较高的权重意味着该股票在指数中的影响力较大,其价格走势与指数的相关性也可能较高。

2. 风格轮动选股。该策略着重于市场风格转换的监测。沪深 300 收盘指数的动态,能够有效映射大盘股的整体表现趋向。当沪深 300 指数呈现上升态势时,这可能预示着市场风格正倾向于大盘股,此时量化投资策略会适时调整,加大对大盘股风格的配置比重,进而对沪深 300 成分股中不同风格股票的筛选产生影响。若量化模型判定当前市场环境适宜大盘股表现,那么该策略将更侧重于选取沪深 300 成分股中市值规模较大、与指数关联度较高的股票。

(二) 择时策略

1. 趋势跟踪。量化投资的趋势跟踪策略会密切关注沪深 300 收盘指数的走势。如果沪深 300 指数呈现出明显的上涨或下跌趋势,量化模型会根据设定的趋势判断规则发出买入或卖出的信号。当沪深 300 指数连续上涨一段时间,突破了某个关键的技术指标或均线系统,量化策略可能会认为市场处于上升趋势,从而买入沪深 300 成分股中的相关股票。反之,如果指数连续下跌,触发了止损或卖出条件,量化策略会卖出股票以规避风险。

2. 均值回归。基于均值回归理论的量化策略认为,股票价格会围绕其均值波动。沪深 300 收盘指数的波动可以帮助量化模型判断成分股价格是否偏离了其合理的价值区间。当沪深 300 指数上涨过高,使得成分股的估值偏高时,量化策略可能会预期股价将回归均值,从而卖出相关股票。当指数下跌过多,成分股估值偏低时,量化策略会买入股票。通过对沪深 300 指数的监测和分析,量化模型可以更准确地把握成分股的买卖时机^[4]。

(三) 套利策略

1. 统计套利。统计套利策略通过挖掘历史数据中的统计规律,寻找价格偏离的机会进行套利。沪深 300 收盘指数与成分股的收盘价之间的差异可以为统计套利提供参考。如果某些成分股的收盘价相对于沪深 300 指数的表现出现异常波动,偏离了历史上的统计关系,量化模型可能会认为存在套利机会。通过同时选择买入价值相对被低估的股票,同时卖出价值

相对被高估的股票,在价格回归正常水平时获得套利收益。

2. 股指期货套利。沪深 300 股指期货与沪深 300 指数之间存在着紧密的关联性,然而,股指期货的价格并非完全受沪深 300 指数支配,其还会受到市场预期、资金状况等多种外部因素的共同作用,可能导致股指期货价格与沪深 300 指数产生一定程度的偏差。量化投资策略正是基于这种价格偏差来进行套利的。具体而言,当沪深 300 股指期货的市场价格高于其基于沪深 300 指数计算所得的理论价格时,投资者可以实施套利策略,即卖出股指期货并同时买入沪深 300 指数的成分股。相反,如果沪深 300 股指期货的市场价格低于其理论价格,投资者则可以选择买入股指期货并卖出沪深 300 指数的成分股,从而实现套利。沪深 300 收盘指数的变化,将进一步影响股指期货与现货之间的价差关系,从而对量化套利策略的实施产生相应的影响^[5]。

四、小结

随着期货市场的持续发展以及不断走向完善,在这个过程中,有一部分传统投资者对于量化投资模式并没有一个较为充分且深入的了解。这种知识层面的欠缺,导致传统投资者在市场中面临着越来越激烈的竞争态势。因为量化投资模式的存在,在交易策略、交易效率等诸多方面都可能对传统投资模式产生冲击,进而使得传统投资者能够获取利润的机会一天天地减少,他们在市场中所面临的盈利难度也在不断增大。不过,另一个角度进行分析,这种情况实际上为量化投资模式投资者在注资环境的优化方面带来了全新的机遇。并且,为确保量化投资在期货市场中能够充分发挥其应用价值,应当正视中国国内市场复杂多变的情况,吸纳并投资更多的人力、物力在中国期货市场中,从而发掘中国量化基金的投资力。◆

参考文献:

- [1] 许龙. 量化投资的应用现状与发展前景[J]. 现代商业, 2019(05): 68-69.
- [2] 张晓燕, 张远远. 量化投资在中国的发展及影响分析[J]. 清华金融评论, 2022(01): 44-45.
- [3] 李占军. 量化投资在我国投资市场中的应用[J]. 投资与合作, 2022(11): 22-24.
- [4] 李丽佼(Giselle Li). 多因子量化投资策略的构建及应用[D]. 电子科技大学, 2022.
- [5] 曾青云. 资本市场量化投资的风险及风险控制策略研究[J]. 投资与合作, 2022(06): 4-6.

(作者简介: 张达, 中煤科工集团金融租赁有限公司助理经济师。)