

<b>From:</b>		<b>To:</b>	
<b>Date:</b>	2025/04/17	<b>Class:</b>	
<b>Title:</b>	AJD001001	<b>Version:</b>	1.0.0

**摘要** 本文围绕豆粕(M)与菜粕(RM)期货品种构建基于协整关系的配对交易策略。通过价格走势、收益率相关性及 ADF 检验,验证二者具备稳定的长期均衡关系,并以滚动 Z 值作为交易信号,引入窗口期  $X$  与阈值  $K$  进行参数优化与敏感性分析。

基础策略在 2014–2021 年表现稳健,但 2022 年后出现回撤放大,回测显示最大回撤超 20%。将止损比例由 5% 下调至 2% 后,策略回撤与波动率显著下降,稳定性提升。

为贴近实盘执行,进一步改用真实收盘价,并在换月日  $T$  日直接平仓,提升模型有效性。针对策略失效时段,局部协整分析发现残差不再平稳,协整关系弱化。建议实盘中引入滚动协整检验、Kalman 滤波与基本面因子增强策略鲁棒性。

**关键词** 跨市场套利, 豆粕期货, 菜粕期货

## 1 策略主要思想

本策略研究基于国内商品期货市场中的两个重要油粕品种——豆粕(M)与菜粕(RM),构建统计套利策略。两者在应用端均为重要的蛋白类饲料原料,需求具有高度替代性,价格走势也呈现出长期同涨同跌的特征。因此,其价差具备协整关系的潜力,适合作为配对交易的对象。

在现实市场中,豆粕与菜粕的价格会受到不同因素的影响。例如,豆粕价格较多受制于美国大豆的进出口、南美大豆供给与压榨环节的边际变化,而菜粕更多受国内油菜籽压榨情况及植物油供需关系的扰动。虽然短期内价格可能出现波动与分化,但在长期内,其价格结构往往维持一定的均衡状态。策略即基于此构造统计套利机会。

本策略的核心思想如下:

- **利用长期协整价差构建配对交易:** 通过建立豆粕与菜粕对数价格的协整关系,提取出两者的残差项作为均衡偏离的衡量标准,判断是否存在套利机会。
- **盈利逻辑来源于短期偏离长期均衡后的均值回复:** 当豆粕与菜粕价格之间出现短期的系统性偏离时,通过在高估资产上做空、低估资产上做多的方式建立双边头寸,待其价差回归均衡后平仓获利。
- **信号源于协整残差的 Z 值:** 将协整残差进行标准化处理,计算其 Z-score; 当 Z 值偏离历史均值超过一定阈值  $K$ , 则视为套利信号生成; 回归至均值区间或达到止盈止损线则进行平仓。

本策略通过对协整残差进行滚动窗口处理,引入观察窗口长度  $X$  与标准差倍数阈值  $K$  作为核心参数,从而动态捕捉套利时机。通过对这两个参数进行系统性遍历和敏感性分析,进一步优化策略表现并控制风险。

## 2 指标分析与信号逻辑

统计套利的核心在于挖掘两种资产之间的价格联动关系,并在其偏离长期均衡时进行双边交易以获取收益。策略能否稳定盈利依赖于以下几个前提条件:

- (1) 两资产的价格序列需具有较强的相关性,即使在市场波动行情下仍能保持一定程度的同步变动;
- (2) 二者需具有长期结构性的联动性,即同涨同跌的价格关系需在较长时期内稳定存在;
- (3) 资产间需存在稳定的协整关系,确保价差具备均值回归的统计基础。

因此,在进入信号构建与策略执行环节前,首先系统检验豆粕与菜粕价格之间是否具备稳定且显著的协整关系。

### 2.1 线性相关性

为了验证豆粕(M)与菜粕(RM)之间是否具备长期线性关系,首先绘制两者的复权收盘价走势,如图1所示。

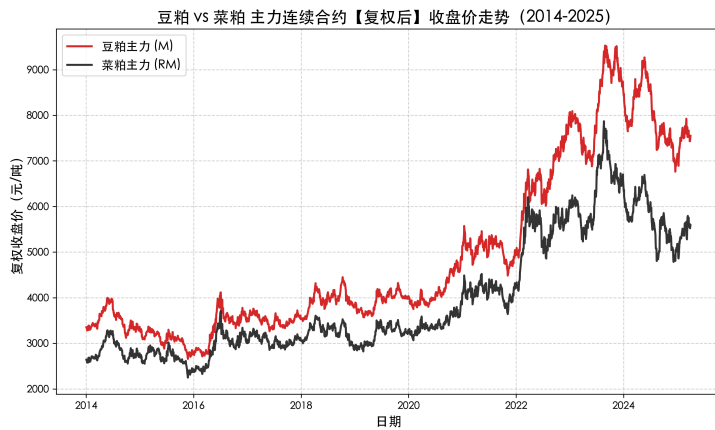


图1 豆粕 vs 菜粕主力连续合约【复权后】收盘价走势(2014-2025)

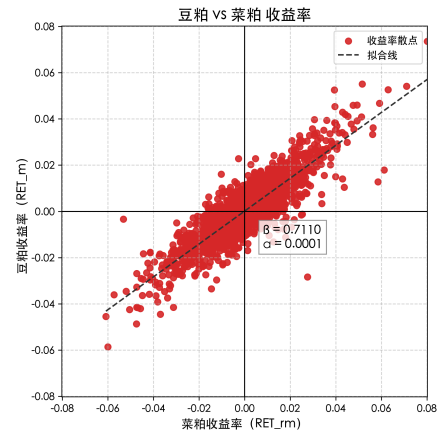


图2 豆粕 vs 菜粕收益率散点图

从图中可以观察到,两者价格走势在大多数时间段表现出明显的协同特征,即同涨同跌。在2019年之前,二者整体波动节奏高度一致;而2022年之后,虽存在一定价差扩张现象,但仍大体维持联动状态。这为后续研究二者是否存在协整关系提供了数据基础和直觉支撑。

进一步地,计算豆粕与菜粕的对数收益率,并绘制了散点图(图2)。从图中可以明显看出,两者收益率之间呈现出高度的正相关关系。

在图中,通过线性回归拟合的斜率( $\beta \approx 0.711$ )和截距( $\alpha \approx 0.0001$ )说明二者收益序列高度正相关,有利于后续构建协整模型。

### 2.2 数据平稳性检验

为了进一步判断豆粕与菜粕价格及其收益率是否满足协整分析的前提条件,对它们的对数价格序列( $\ln P$ )与对数收益率序列( $r_t = \ln P_t - \ln P_{t-1}$ )分别进行了单位根检验(ADF检验)与格兰杰因果

性检验(Granger Causality Test)。

### 2.2.1 ADF 平稳性检验

表 1 展示了对豆粕与菜粕的对数价格和收益率序列的 ADF 检验结果。

表 1 ADF 平稳性检验结果

变量	ADF 统计量	p 值	临界值 1%	临界值 5%	临界值 10%	结论
$\ln P_m$	-0.5329	0.8854	-3.4327	-2.8626	-2.5673	非平稳(不拒绝单位根)
$\ln P_{rm}$	-1.1437	0.6973	-3.4327	-2.8626	-2.5673	非平稳(不拒绝单位根)
$RET_m$	-31.6906	0.0000	-3.4327	-2.8626	-2.5673	平稳(拒绝单位根)
$RET_{rm}$	-31.8242	0.0000	-3.4327	-2.8626	-2.5673	平稳(拒绝单位根)

可以看出,豆粕与菜粕的对数价格序列均存在单位根,表现为非平稳过程;而其对数收益率序列在 1% 显著性水平下显著平稳。此结果符合金融资产常见的“价格非平稳、收益率平稳”的特征,也为构建协整模型奠定了必要的统计基础。

### 2.2.2 Granger 因果关系

为了进一步验证豆粕与菜粕之间是否存在动态联动关系,对其平稳的收益率序列进行 Granger 因果性检验,结果如表 2 所示。

表 2 Granger 因果检验结果(最大滞后阶数 = 5)

滞后 阶数	$RET_{rm} \rightarrow RET_m$			$RET_m \rightarrow RET_{rm}$		
	F 统计量	p 值	显著性	F 统计量	p 值	显著性
1	7.0646	0.0079	***	3.7689	0.0523	*
2	3.5471	0.0289	**	2.5041	0.0819	*
3	2.3287	0.0726	*	1.7529	0.1541	-
4	1.8434	0.1178	-	1.3156	0.2617	-
5	1.5794	0.1624	-	1.3164	0.2540	-

注:\* 表示 10% 显著性水平( $p < 0.10$ ), \*\* 表示 5% 显著性水平( $p < 0.05$ ), \*\*\* 表示 1% 显著性水平( $p < 0.01$ )。

从表中可以看出,在较小的滞后阶数下(1 阶和 2 阶),两个方向均存在 Granger 因果关系,尤其是  $RET_{rm} \rightarrow RET_m$  的方向在 1% 水平下显著。说明菜粕的价格变动在一定程度上可用于预测豆粕价格的变动,二者之间具有统计上的联动性。

综合 ADF 与 Granger 检验结果,可以认为豆粕与菜粕收益率序列平稳,且存在显著的动态因果关系,符合构建协整模型的前置条件。

## 2.3 协整模型构建与检验

### 2.3.1 OLS 协整回归模型

为构造豆粕与菜粕价格之间的协整关系,选取其对数价格序列  $\ln P_m$  和  $\ln P_{rm}$ , 采用最小二乘法(OLS)构建以下线性回归模型:

$$\ln P_m = \alpha + \beta \ln P_{rm} + \varepsilon_t$$

其中  $\varepsilon_t$  为回归残差项,后续将作为套利信号的基础。图 3 展示了实际  $\ln P_m$ , 拟合  $\ln \hat{P}_m$  与残差项

的变化趋势。

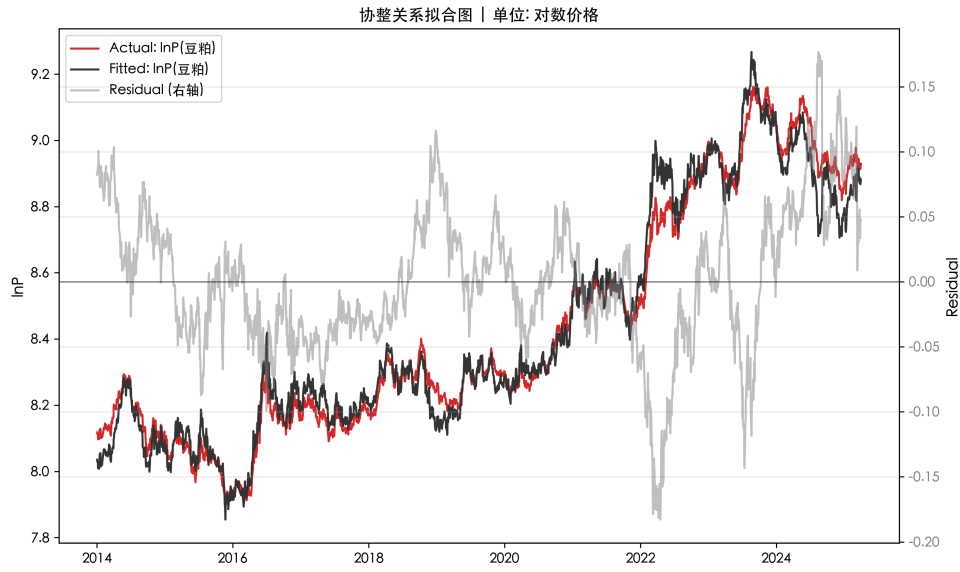


图3 协整关系拟合图:对数价格与残差序列

根据 OLS 回归结果:  $R^2 = 0.976$ , 表明模型能解释 97.6% 的价格波动; 协整回归系数  $\beta = 1.1278$ , 显著为正; 残差在图中呈现出均值回归特性。

$$\ln P_m = -0.8499 + 1.1278 \cdot \ln P_{rm} + \varepsilon_t$$

其中  $\varepsilon_t$  为残差项, 作为套利信号的主要来源。根据协整关系式的  $\beta$ , 菜粕价格上涨 1%, 意味着豆粕上涨 1.1278%。若要进行双边交易, 控制波动率相等并且排除小数的影响, 豆粕做 8 手, 菜粕做 9 手。

### 2.3.2 协整结构的稳健性与不对称性检验

为检验豆粕与菜粕价格关系在不同市场状态下是否存在结构性变化, 进一步引入菜粕涨跌方向的虚拟变量  $D_{rm}$ , 并构造交互项  $D_{rm} \cdot \ln P_{rm}$ , 用于刻画上涨与下跌状态下的协整斜率差异。

$$\ln P_m = \alpha + \beta_1 \ln P_{rm} + \beta_2 D_{rm} + \beta_3 (D_{rm} \cdot \ln P_{rm}) + \varepsilon_t$$

- $D_{rm} = \mathbb{I}(\Delta P_{rm} > 0)$ , 表示当日菜粕价格是否上涨;
- $\beta_1$  表示在菜粕下跌状态下的协整斜率;
- $\beta_1 + \beta_3$  表示在菜粕上涨状态下的协整斜率。

OLS 回归  $R^2 = 0.976$ , 与原协整模型一致;  $\beta_1 = 1.1363$ , 和原值相差不大; 交互项系数  $\beta_3 = -0.0163$ , 在 5% 显著性水平下显著。

虽然从交互项统计显著性上看, 协整关系存在微弱的结构差异, 但其数值较小, 经济含义不强, 策略执行中并不采用不对称手数配比, 仍使用豆粕 8 手、菜粕 9 手的固定配比方案。

### 2.3.3 残差 ADF 检验

为验证  $\varepsilon_t$  是否为平稳过程, 对残差进行了 ADF 单位根检验。检验结果如下:

- ADF 统计量: -3.7097, p 值: 0.0040

由于 ADF 统计量小于临界值, 且  $p$  值远小于 0.05, 因此拒绝原假设(残差具有单位根), 残差序列平稳。由此可判定豆粕与菜粕价格序列间存在显著协整关系, 符合构造统计套利策略的关键条件。

## 2.4 信号生成逻辑

在得到协整残差序列后, 采用标准化残差  $Z$ -score 作为交易信号的生成依据。其定义如下:

$$Z_t = \frac{\varepsilon_t - \mu_t^{(X)}}{\sigma_t^{(X)}}$$

其中  $\mu_t^{(X)}$  与  $\sigma_t^{(X)}$  分别表示残差在过去  $X$  个交易日的滚动均值与标准差。

当  $Z$  值出现极端偏离时, 意味着豆粕与菜粕的价格关系偏离长期均衡, 此时生成套利信号。具体规则如下:

- 若  $Z_t > K$ , 说明价差显著偏高, 策略做空豆粕、做多菜粕;
- 若  $Z_t < -K$ , 说明价差显著偏低, 策略做多豆粕、做空菜粕;

当开仓信号设置得越“宽松”, 策略能够捕捉到更多潜在的套利机会, 但随之而来的是更频繁的交易和更高的交易成本; 而若信号条件设置得过于“严苛”, 虽然能有效控制成本, 却可能错失部分有利的交易时机。因此, 需要控制  $X$  和  $K$ , 在机会与成本控制之间取得平衡。

## 2.5 交易设定

- (1) **信号生成时点**: 每个交易日收盘后, 根据协整残差计算当日的标准化  $Z$  值;
- (2) **开仓时间**: 若  $T$  日  $Z_t > K$  或  $Z_t < -K$ , 则在下一交易日 ( $T+1$ ) 以当日收盘价 **建仓**;
- (3) **平仓时间**: 当浮盈浮亏超过阈值, 或残差回归至均值附近 ( $|\varepsilon_t - \mu_t| < 0.005$ ) 时,  $T$  日生成平仓信号, 在  $T+1$  日执行 **双边平仓**;
- (4) **仓位设置**: 固定手数, 豆粕做 8 手, 菜粕做 9 手, 合约乘数均为 10;
- (5) **资金管理**: 初始资金设定为 200,000 元, 浮盈浮亏按日更新;
- (6) **交易价格**: 统一使用收盘价成交, 方向性滑点设定为  $\pm 0.01\%$ ;
- (7) **止盈止损设置**: 基础止盈/止损设定为 5%;
- (8) **信号参数遍历范围**:  $K \in \{0.2, 0.3, \dots, 2.0\}$ ,  $X \in \{2, 3, 5, \dots, 100\}$ ;
- (9) **数据周期**: 2014 年 1 月起至 2025 年 3 月, 数据频率为日度(主力连续合约, 复权);

### 3 基础策略回测结果

#### 3.1 基础回测表现

在初始策略中, 设置止盈/止损阈值为 **5%**, 采用对齐最开始复权乘数为 1 的复权收盘价进行估值, 交易价格为方向性滑点后的次日收盘价。通过遍历参数  $X \in [2, 100]$ ,  $K \in [0.2, 2.0]$ , 寻找最佳参数组合。以下为回测核心图示。

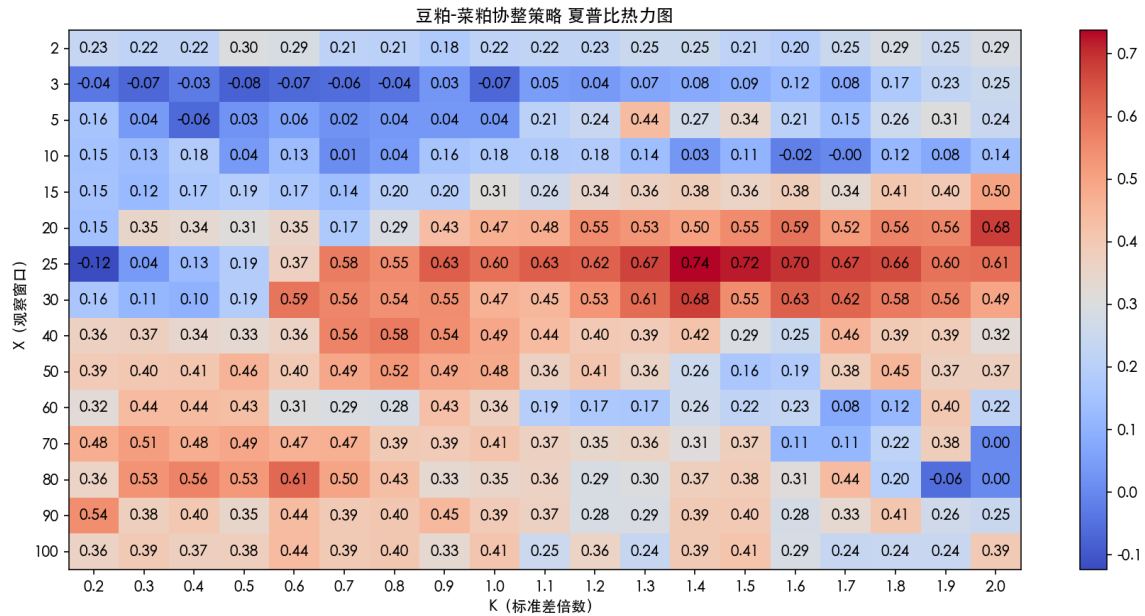


图4 Sharpe 比率热力图(止损 = 5%, 复权价格)

图4 表示不同参数组合  $(X, K)$  下的夏普比率分布, 最优组合为  $X = 25, K = 1.4$ , 对应最大 Sharpe 比率为 **0.74**。

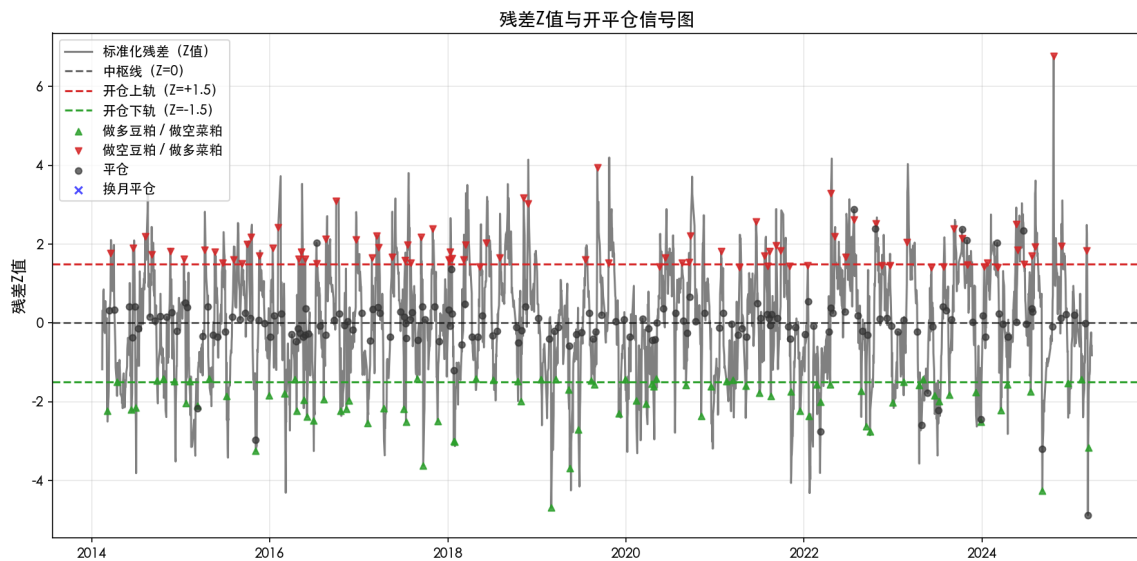


图5 残差 Z 值与开平仓信号图(最佳参数:  $X = 25, K = 1.4$ )

图5 展示了残差 Z 值随时间的波动及其对应的开仓、平仓等交易信号。其中红色倒三角表示做空豆粕/做多菜粕, 绿色正三角为做多豆粕/做空菜粕, 灰色圆点为平仓, 策略对残差偏离反应灵敏。

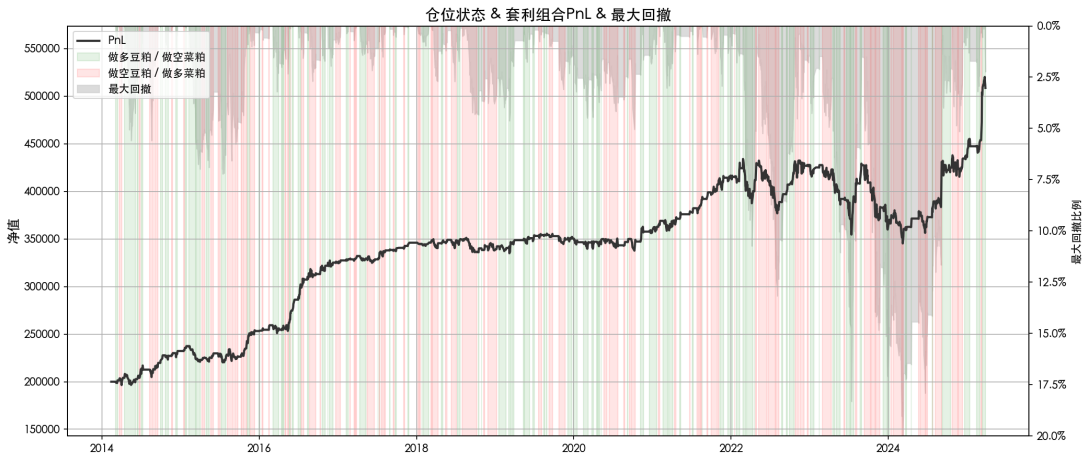


图 6 净值曲线与仓位状态图(止损 = 5%,  $X = 25$ ,  $K = 1.4$ )

图 6 显示策略净值曲线表现稳定，在 2014–2022 年间稳步上升，但在 2022–2024 年出现较大 drawdown(最大回撤超过 20%)。指标如下：

年化收益率	波动率	Sharpe 比率	Calmar 比率	最大回撤	回撤修复时间
9.08%	10.52%	0.737	0.444	20.45%	230 天

表 3 策略回测指标

该策略在平稳阶段表现良好,但在市场结构变化期间,出现较长时间的回撤未能修复。

3.2 参数敏感性分析

图 4 所示 Sharpe 热力图表明：

- $X \in [20, 60]$  区间稳定性较强；
- $K \in [1.2, 1.6]$  处表现较优,表明适中阈值能权衡信号捕捉与交易频率；
- 太小的  $X$  或  $K$  容易导致过拟合与过度交易,Sharpe 下降明显；

热力图形状呈现“热带弧形”,反映参数有一定鲁棒性,但在临界区( $K < 0.6$ 、 $X < 10$ )易造成策略失效。

### 3.3 策略调整实验

针对回撤过大、回撤修复时间长等问题,对基础策略进行调整:止损从 5% 下调为 2%,提高止损敏感性。

调整后重新遍历参数,最佳结果出现在  $X = 20, K = 2.0$ ,如下图所示:

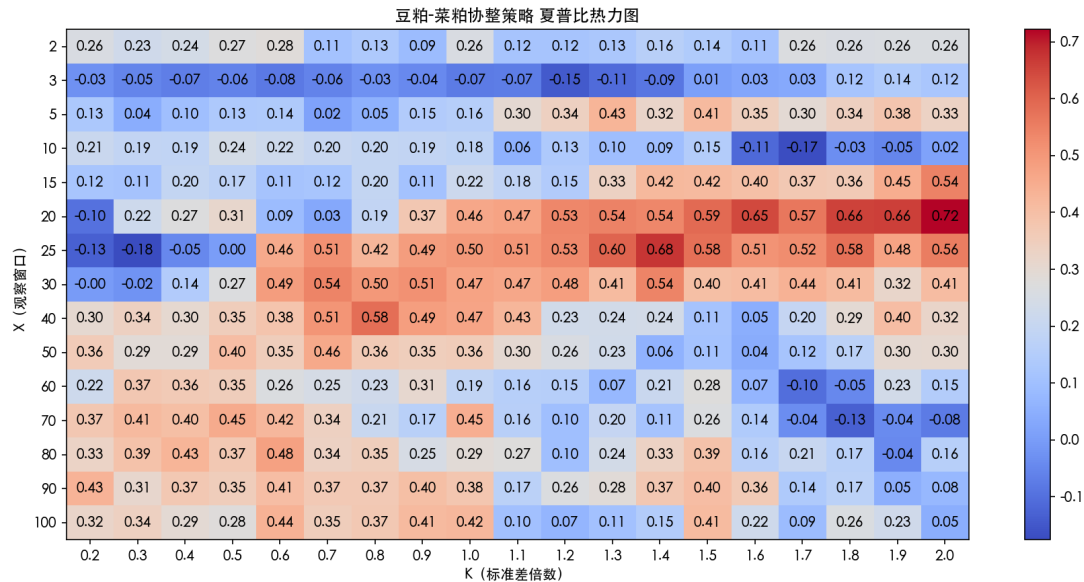


图7 Sharpe 比率热力图(止损 = 2%, 复权价格)

图7可见:在较大开仓阈值下,Sharpe 比率集中分布于  $K \geq 1.5$  区域,说明信号需更强偏离才具备高效表现,反映策略对“极端价差”的敏感度提升。

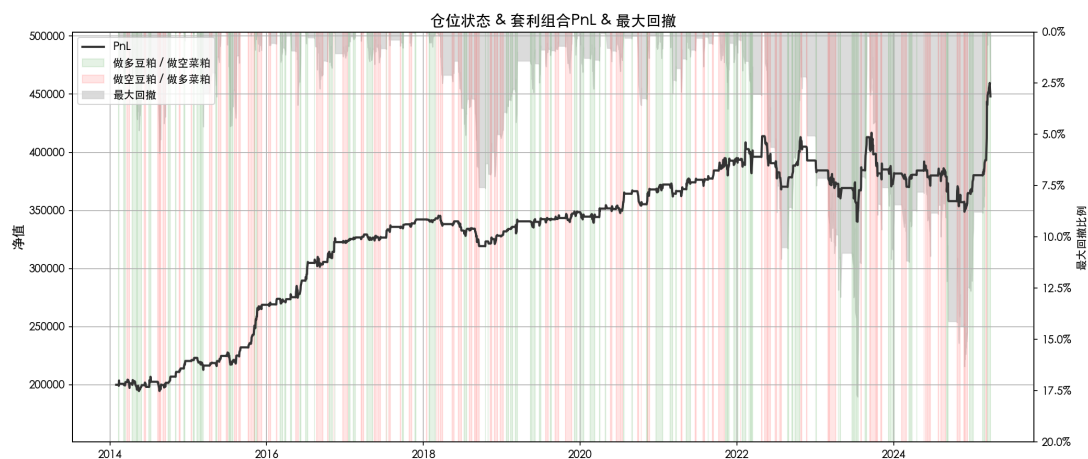


图8 净值曲线与仓位状态图(止损 = 2%,  $X = 20, K = 2.0$ )

图8显示策略在经历换月处理与止损增强后,净值增长更为平滑,最大回撤下降至 17.78%,且回撤修复时间从 230 天下降至 64 天,策略抗风险能力得到明显增强。

年化收益率	波动率	Sharpe 比率	Calmar 比率	最大回撤	回撤修复时间
7.79%	8.87%	0.722	0.438	17.79%	64 天

表4 调整后策略回测指标



虽然年化收益有所下降,但收益的稳定性与回撤控制能力大幅改善,为进一步构建多因子动态协整策略或结合宏观筛选提供了更稳健的基础。

#### 4 考虑真实价格交易

在前述实验中主要采用了对齐后的复权价格进行策略回测。然而观察发现,豆粕与菜粕的复权乘数随时间显著变化,在 2014 年至 2025 年,从 1 点多提升至 3 点多,导致复权价格与真实价格在近年出现较大偏离(见图9与图10)。这种偏差可能影响交易信号的精确性和资金使用效率。

因此,本节在策略回测中改为使用未复权的真实价格,并针对换月日的交易处理方式进行了修正:当合约面临换月时,T 日(即当日)便直接执行双边平仓操作,避免 T+1 价差异异常带来的影响。平仓时也考虑滑点和手续费为 0.01%。



图9 菜粕主力合约未复权 vs 复权收盘价走势



图10 豆粕主力合约未复权 vs 复权收盘价走势

#### 4.1 止盈止损 5% 回测结果

首先在止盈止损比例为 5% 的基础下进行回测,并遍历参数  $(X, K)$  组合,最终最优参数为  $X = 25$ ,  $K = 1.5$ ,其对应的回测指标如下:

年化收益率	波动率	Sharpe 比率	Calmar 比率	最大回撤	回撤修复时间
5.91%	6.77%	0.662	0.249	23.70%	—

表5 真实价格 + 5% 止盈止损 ( $X=25, K=1.5$ )

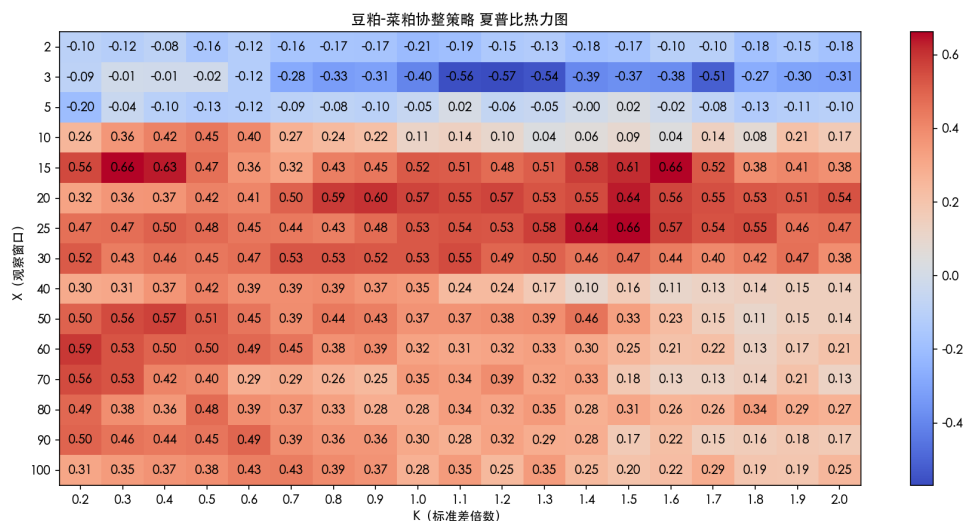


图11 豆粕-菜粕协整策略夏普比热力图(真实价格 + 5% 止盈止损)

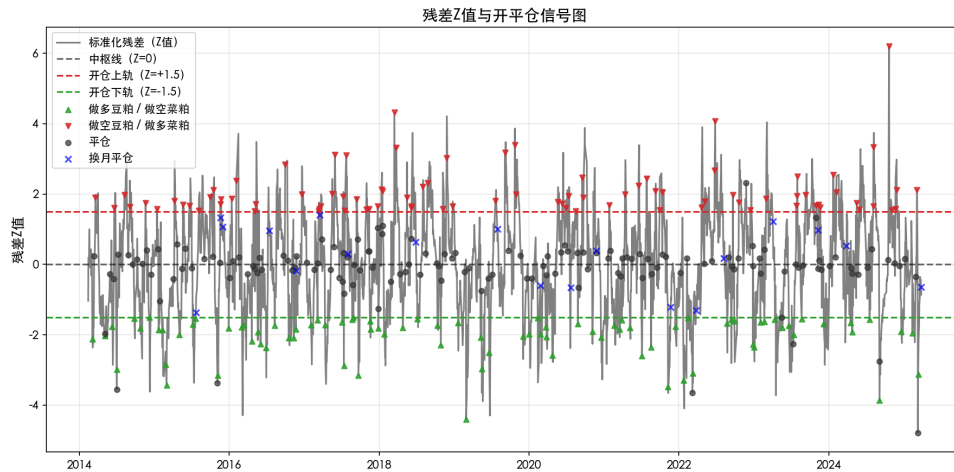


图 12 残差 Z 值与开平仓信号图(真实价格 + 5% 止盈止损)

图中可见, Z 值在偏离阈值  $K = 1.5$  时清晰触发开仓信号, 策略响应及时, 平仓位置亦较为合理。并且在图中用蓝色叉号标明了换月平仓的时间。

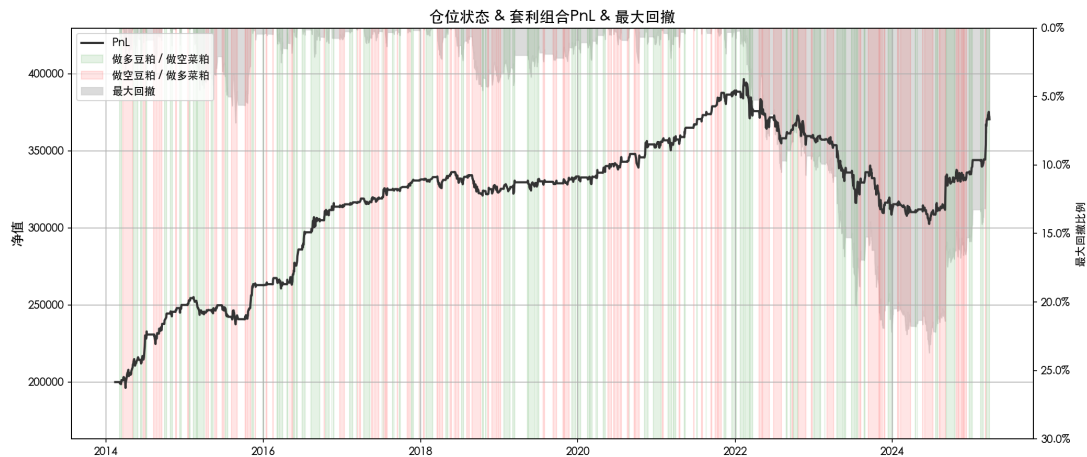


图 13 仓位状态 套利组合 PnL 最大回撤(真实价格 + 5% 止盈止损)

相较于使用复权价格的基础策略, 真实价格下策略收益略有下降, 且目前还未修复最大回撤, 但整体表现更加贴近实际交易情况, 回撤结构与信号分布更为稳定, 具备更强的实盘适应性。

#### 4.2 进一步收紧止损至 2%

随后, 将止损比例收紧至 2%, 保持 5% 止盈不变, 重新进行回测。最优参数变为  $X = 20$ ,  $K = 2.0$ , 对应回测结果如下:

年化收益率	波动率	Sharpe 比率	Calmar 比率	最大回撤	回撤修复时间
5.31%	6.06%	0.639	0.285	18.63%	—

表 6 真实价格 + 2% 止损 ( $X=20$ ,  $K=2.0$ )

从回测结果可见, 虽然止损更严格在一定程度上降低了策略收益, 但也有效压缩了回撤水平和波动率, 提升了 Calmar 比与稳定性, 体现出在真实价格下进行严格风控管理的重要性。

## 5 风险观察与失效分析

在策略运行过程中,观察到 2022 年至 2024 年间策略表现明显走弱,部分持仓区间出现连续回撤。为此,对协整关系与残差稳定性进行了进一步分析,以识别潜在的策略失效风险。

### 5.1 协整关系波动加剧

回顾图 3 协整拟合图可以发现,整体协整关系虽保持稳定,但 2022 年左右残差波动显著放大,拟合误差增大,说明该阶段两资产之间的长期均衡关系发生了一定程度的偏离。可能的原因包括:

- 豆粕与菜粕产业链在俄乌冲突、能源危机期间受到不同程度冲击;
- 国内外政策对两个品种的影响非同步;
- 供需逻辑在极端行情下发生短期结构性变化。

### 5.2 局部样本协整关系失效

为进一步验证该阶段是否存在协整失效的风险,截取 2022-02-10 至 2024-03-31 这一期间的数据片段,重新进行 OLS 拟合与残差 ADF 检验,结果如下:

- **ADF 统计量:**  $-1.4581$ , **p 值:**  $0.5541$

该检验结果表明,残差序列在该阶段并不平稳,未能拒绝单位根假设。换言之,该时间段内豆粕与菜粕价格的协整关系失效,协整基础动摇,导致策略信号质量显著下降,部分开仓无法实现均值回归。

## 6 总结与展望

### 6.1 当前结论

- 策略在 2014–2021 年期间表现优异,整体收益稳健,具有较好的可操作性;
- 参数稳定性较强,  $(X, K)$  组合的热力图呈现连续分布,体现出交易逻辑的鲁棒性;
- 最大回撤水平偏高,特别是在 2022 年后,策略收益波动显著上升,风险控制有待进一步强化。

### 6.2 改进方向

- 使用 Kalman 滤波等动态线性模型实时估计协整系数,提升模型响应性;
- 引入滚动协整检测机制,结合滑动窗口进行动态参数  $(X, K)$  选择;
- 换月时进行单边换仓而不是直接双边平仓;
- 融入基差、库存、进口成本等基本因子,以增强信号解释力与抗扰性;
- 进一步拓展至油脂产业链多品种套利策略(如豆油、棕榈油、菜油),提升策略广度与稳健性。