

配对交易系列（一） 基于 Copula 函数的配对交易研究

投资咨询业务资格：
证监许可【2012】669号

报告要点

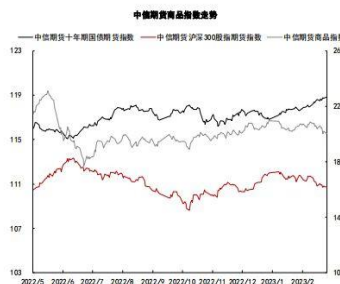
本篇报告利用 Copula 函数提出了两种创新型配对策略：单一 Copula 策略和混合 Copula 策略，对标协整策略在商品和股指期货上分别进行了回测，发现 Copula 函数从配对相关性变化进行考虑，在捕捉行情变化时会更加敏锐；此外，若将具有不同分布特征的阿基米德 Copula 函数进行线性组合，能再次加强策略的普适性。

摘要：

Copula 函数被看成联合分布和各个变量边缘函数之间的桥梁，因此可以用来描述变量之间的相关性，本文根据 Copula 函数的特性，提出了两种创新型配对策略：单一 Copula 策略和混合 Copula 策略。同时，在进行配对选择时，计算出了不同配对的 Pearson、Spearman 和 Kendall 三大相关系数来对配对进行排序和筛选。此外，所有配对选择和模型搭建皆采用轮动机制来增强策略的时效性和多变性。

对标经典协整策略回测后发现，混合 Copula 策略显著优于协整策略，年化收益率达到 16.62%，较协整策略提高了 10% 以上；夏普比率和卡玛比率分别达到 1.22 和 1.72，是协整策略的 2.9 倍和 5.9 倍。混合 Copula 函数较协整策略对相关性的变化具有更强的敏锐性，因此交易轮次和盈亏比都得到了优化；同时，混合 Copula 函数融合了三大阿基米德 Copula 函数不同的分布特点，因此具有更强的普适性，最大回撤比率得到了更有效的控制。

此外，我们发现在三个策略的回测中，化工板块配对的收益情况都更为良好且稳定，因此推测化工板块品种之间具有更强的联动性，在进行配对交易时，可以着重考虑该板块。



商品量化组

研究员：
蒋可欣 FRM
jiangkexin@citicsf.com
从业资格号 F03098078
投资咨询号 Z0018262

时序预测系列研究 报告

专题报告一：基于分解算法和深度学习的预测建模研究——
20230216

专题报告二：如何利用
Transformer-based 模型对价格进行“天气预报”——
20230411

风险提示：本报告中所涉及的算法和模型应用仅为回溯举例，并不构成推荐建议。

目录

摘要:	1
一、引言	3
二、协整法	3
三、Copula 函数	4
(一) 椭圆 Copula	5
(二) 阿基米德 Copula	6
四、配对选择	7
(一) 相关性测量	7
(二) 品种筛选	9
(三) 配对挑选规则	10
五、策略构建和实证分析	10
(一) 回测说明	10
(二) 协整策略	11
(三) 单一 Copula 函数策略	12
(四) 混合 Copula 函数策略	14
六、结论与展望	15
参考文献	17
免责声明	18

图表目录

图表 1: 阿基米德 Copula 函数的生成元和概率分布函数	7
图表 2: 配对品种库	9
图表 3: 回测说明	11
图表 4: 协整策略回测净值变化	11
图表 5: 协整策略回测结果	11
图表 6: 单一 Copula 函数策略回测净值变化	12
图表 7: 单一 Copula 函数策略回测结果	13
图表 8: 混合 Copula 函数策略回测净值变化	14
图表 9: 混合 Copula 函数策略回测结果	14
图表 10: 各策略回测净值变化综合	15
图表 11: 回测结果总结	15

一、引言

配对交易是一种重要的统计套利策略，也是一种市场中性策略。其基本思路比较简单，通过统计的方法寻找两个长期趋势相似的资产进行配对，若两者价差偏离其长期均值，买入被低估的资产，卖出被高估的资产，随后等待价差回归长期均值以此获利。目前研究配对交易主要有三种方法：**最小距离法**、**协整法**和**随机价差法**。在**最小距离法**中通过计算不同配对价格序列之间的平方距离，挑选出距离最小的 N 个配对进行套利，该方法简单易于操作，但对价格偏差的判定停留在静态的标准上，与具有时变性的真实市场相违背，因此仅适用于短期交易。**随机价差法**认为不同资产的价差由白噪音和一个均值回归过程组成，利用状态空间模型对资产 A 、 B 价格序列的差值进行了建模，此模型是一个连续模型，因此可以用来预测。但随机价差法要求长期内配对资产的均值回归水平需相同，这点严重限制了随机价差法的使用范围。**协整法**是目前使用最为广泛的套利策略之一，描述了非平稳时间序列之间存在的长期稳定均衡关系，即两个非平稳时序的线性组合所得序列是平稳的，可以使用 EG 两步法来检测配对间的协整关系。协整法的瓶颈在于其描述的均衡关系特指的是线性关系，而在现实生活中均衡关系存在非对称性，尾部渐进相关或独立模式等情况，仅依靠线性来刻画依赖关系是不够的。

可以发现传统的配对交易方法都存在各自的缺陷，在面对当前越发复杂的金融市场时表现得逐渐吃力。随着各类理论和 AI 技术的蓬勃发展，各类新型技术被尝试运用在配对交易中进行优化，例如**深度学习和强化学习**等技术都取得了比较显著的效果，本系列将聚焦在**如何利用新型理论算法对传统技术进行优化**这一问题上，为大家介绍更多更为灵活且高效的配对交易方法。**Copula 函数**是近年来刻画**非正态、非线性相关结构**的有效方法之一，本篇为本系列的第一篇，将尝试运用 Copula 系列函数来对协整法的不足进行弥补，分为四大部分，第一部分和第二部分将会分别介绍协整法和 Copula 系列函数的基本理论概念和相对应的具体策略构建；第三部分将利用 **Kendall 秩相关系数**对当前市场品种进行配对筛选；第四部分将对基于协整法和 Copula 系列函数的套利策略进行回测对比。

二、协整法

协整法依托于协整理论，于 1987 年被 Engle 和 Granger 提出，针对非平稳序列建立回归模型，若序列的线性组合序列为平稳序列，则认为此非平稳序列

之间存在长期稳定的均衡关系。通常使用 EG 两步法或 Johansen 检验进行协整检验，两者的区别点在于 EG 两步法采用的是一元方程技术，而 Johansen 检验是采用了多元方程技术，使用范围会更广。本篇研究聚焦在两个品种配对，使用 EG 两步法会更加简单便捷，因此本文主要讨论和介绍 EG 两步法的检验步骤，具体步骤如下：

1. 对两个合约的价格序列 (P_1, P_2) 取对数，得到序列 $(\ln P_1, \ln P_2)$ ；
2. 对序列 $(\ln P_1, \ln P_2)$ 分别进行平稳性检验，如若不平稳，进行差分和再次平稳性检验，直到序列平稳为止，检验所得的两个变量序列必须是相同阶数的单整过程才有存在协整关系的可能；
3. 对符合要求的序列进行 OLS 建模得到以下方程：

$$\ln P_2 = \alpha + \beta \ln P_1 + resid \text{ (公式 1)}$$

4. 对公式中的残差项 $resid$ 进行平稳性检验，若残差序列为平稳序列，则表明两个合约存在协整关系。

若两个合约中存在协整关系，公式 1 中的 β 则为两个合约的协整系数，即套利交易中的买卖比例。此时，若协整回归后的残差项超出一定水平，则认为市场出现异常，进行相应的买卖操作，当残差项回归正常水平时再进行相应的平仓操作。使用指标 $Sign1$ 作为残差偏移水平的判定规则，具体计算公式如下：

$$sign1_t = resid_t / std(resid) \text{ (公式 2)}$$

其中， $resid_t$ 为使用前 N 个交易日数据回归后得出的当前残差项， $std(resid)$ 为前 N 个交易日残差项的标准差， $sign1_t$ 为当前信号。当 $sign1_t$ 绝对值大于 1 时，认为协整关系发生了一个较强的偏移；当 $sign1_t$ 的绝对值大于 2 时，认为协整关系不再存在。因此设定相应的交易规则：当 $sign1_t$ 的绝对值超过 1 且回落到 1 时开仓，若此时 $sign1_t$ 大于 0，买 P1 卖 P2，若此时 $sign1_t$ 小于 0，卖 P1 买 P2；开仓后，当 $sign1_t$ 的绝对值回落到 0 时平仓，或当绝对值超过 2 时止损平仓。

三、Copula 函数

Copula 函数中的 Copula 来自于拉丁语，意为“连接”，因此又被称为连接

函数。Copula 函数最早由 Sklar 在 1959 年提出，Sklar 认为任何一个联合分布函数都可以被分解为多个边缘分布函数和一个 Copula 函数，因此 Copula 函数可以被看成联合分布和各个变量边缘函数之间的桥梁，用来描述变量之间的相关性，具体公式如下：

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)) \text{ (公式 3)}$$

其中，C 代表了 Copula 函数， $H(x, y)$ 为一个具有连续边缘分布的 $F(x)$ 和 $G(y)$ 的二元联合分布函数， $F(x)$ 和 $G(y)$ 分别为两个任意的概率分布函数。 $F^{-1}(u)$ 为 $F(u)$ 函数的反函数，又被称为 CDF 的逆变换或逆累积分布函数。虽然 Copula 函数的形式非常多，但每个 copula 函数都必须满足以下几点基本性质：

1. 定义域为 $[0, 1] * [0, 1]$ ，值域为 $[0, 1]$ ，即 $C: [0, 1] * [0, 1] \rightarrow [0, 1]$ ；
2. 对任意一个自变量，C 均为单调非减函数；
3. 对任意 $u, v \in [0, 1]$ ，满足： $C[u, 0] = C(0, v) = 0$ ， $C(u, 1) = C(1, u) = u$ ， $C(v, 1) = C(1, v) = v$ ；
4. 对于任意 $0 \leq u_1 \leq u_2 \leq 1$ 和 $0 \leq v_1 \leq v_2 \leq 1$ ，满足 $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$ 。

目前，Copula 函数主要分为两大类，分别为：椭圆 Copula 函数簇 (Elliptic Copula) 和阿基米德 Copula 函数簇 (Archimedean Copula)，我们——来看。

(一) 椭圆 Copula

椭圆 Copula 类函数的最大特点就是其分布图形呈椭圆形状，典型的椭圆 Copula 函数有 高斯 Copula 函数和 t-Copula 函数，两者均有对称的尾部相关性，主要区别点体现在尾部的厚度，我们分开来看。

1. 高斯 Copula 函数

高斯 Copula 函数又被称为正态 Copula 函数，就是随机向量的各个边缘分布和连接各边缘分布的 Copula 函数都是正态分布的情况，其二元概率分布函数如公式四所示：

$$C_{Gaussian}(u, v; \rho) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \frac{-(r^2 + s^2 - 2\rho rs)}{2(1-\rho^2)} dr ds \quad (\text{公式 4})$$

其中, $\rho \in (-1, 1)$ 为相关系数, $r = \Phi^{-1}(u)$, $s = \Phi^{-1}(v)$, Φ^{-1} 为标准正态分布函数的逆函数。

2. t-Copula 函数

t-Copula 函数中随机变量服从自由度为 v 的 t 分布, 其二元分布函数可以写成如下形式:

$$C_{St}(u_1, u_2; \rho, v) = \int_{-\infty}^{t_v^{-1}(u_1)} \int_{-\infty}^{t_v^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left[1 + \frac{s^2 + t^2 - 2\rho st}{v(1-\rho^2)} \right]^{-\frac{v+2}{2}} ds dt \quad (\text{公式 5})$$

其中, $\rho \in (-1, 1)$ 为相关系数, $s = t_v^{-1}(u_1)$, $t = t_v^{-1}(u_2)$, v 为自由度。

(二) 阿基米德 Copula

阿基米德 Copula 函数簇的分布函数定义最早由 Genest 和 Mackay 在 1986 年提出, 和椭圆 Copula 函数相比, 它的计算和构造都简便很多, 具体公式如下:

$$C(u_1, u_2, \dots, u_N) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2) + \dots + \varphi(u_N)) \quad (\text{公式 6})$$

其中, 函数 $\varphi(\cdot)$ 为阿基米德 Copula 函数的生成元, 也叫生成器(generator function), 必须满足以下条件:

1. $\varphi: [0, 1] \rightarrow [0, \infty]$, 对于 $t \in (0, 1)$, $\varphi'(t) < 0$, $\varphi''(t) \geq 0$, 即 φ 是连续的减凸函数, 且 $\varphi(1) = 0$;
2. $\varphi^{-1}: [0, \infty] \rightarrow [0, 1]$ 是完全单调函数,

$$\varphi^{-1}(t; \theta) = \begin{cases} \varphi^{-1}(t; \theta), & 0 \leq t \leq \varphi(0, \theta) \\ 0, & \varphi(0, \theta) \leq t \leq \infty \end{cases}.$$

根据不同的生成元函数可以构造不同的阿基米德 Copula 函数, 常见的

Copula 函数有 Clayton Copula、Gumbel Copula 和 Frank Copula 函数，它们各自具有不同的特色，适合运用在相应的场景中，具体生成元函数和概率分布函数见图表 1。

图表 1：阿基米德 Copula 函数的生成元和概率分布函数

Copula 函数名	概率分布函数	生成元函数
Clayton Copula	$C_{Clayton}(u, v; \theta) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \theta > 0$	$\frac{1}{\theta}(t^{-\theta} - 1)$
Gumbel Copula	$C_{Gumbel}(u, v; \theta) = \exp\left(-[(-\ln u)^\theta + (-\ln v)^\theta]^{\frac{1}{\theta}}\right), \theta \geq 1$	$(-\ln t)^\theta$
Frank Copula	$C_{Frank}(u, v; \theta) = -\frac{1}{\theta} \ln \left[1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{\exp(-\theta) - 1} \right], \theta \in \mathbb{R} \setminus \{0\}$	$-\ln \left(\frac{\exp(-\theta t) - 1}{\exp(-\theta) - 1} \right)$

资料来源：中信期货研究所

这三种函数由于生成元不同，具有不同的概率分布特征。Clayton 函数的密度函数呈“L”字型，因此适用于描述下尾明显而上尾不明显的相关性；Gumbel 函数的特点正好与 Clayton 函数相反，更适用于描述上尾明显而下尾不明显的相关性；Frank 函数的密度函数呈“U”字型，因此更适用于描述具有对称结构的相关性。阿基米德 Copula 函数和椭圆 Copula 函数相比更为灵活，且更适用于现实生活中非对称和非线性的情况。但值得注意的是当前金融市场越来越复杂，仅凭单一 Copula 函数难以刻画变量之间多变的相关性，因此提出了混合 Copula 的概念。如公式 7 所示，通过线性组合将 Clayton Copula、Gumbel Copula 和 Frank Copula 函数组合成全新的 Copula 函数，其中， $0 \leq \lambda_1, \lambda_2, \lambda_3 \leq 1, \lambda_1 + \lambda_2 + \lambda_3 = 1, \Phi = (\lambda_1, \theta; \lambda_2, \alpha; \lambda_3, \beta)$ 。

$$C_{CGF}(u, v; \Phi) = \lambda_1 C_{Clayton}(u, v; \theta) + \lambda_2 C_{Gumbel}(u, v; \alpha) + \lambda_3 C_{Frank}(u, v; \beta) \text{ (公式 7)}$$

混合 Copula 函数综合了单一阿基米德 Copula 函数的各自特点，通过权重的调整能更有效地刻画不同情况下的相依关系。

四、配对选择

（一）相关性测量

相关系数可以体现两个变量之间变化的趋势方向和趋势程度，因此本文依据相关系数对品种配对组合进行先一步的筛选。目前，统计学上主要有三大相

关系数，分别为：皮尔逊相关系数 (Pearson)、斯皮尔曼秩相关系数 (Spearman) 和肯德尔秩相关系数 (Kendall)，它们适用于不同场景，我们先来依次看一下。

1. Pearson 相关系数

Pearson 相关系数是最常见的相关系数，适用于呈正态分布的连续变量，具体计算公式如下：

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \quad (\text{公式 8})$$

其中，E 表示数学期望。Pearson 相关系数虽然计算简单易于理解，但只能用于描述线性关系。

2. Spearman 秩相关系数

“秩”是 Spearman 和 Kendall 相关系数的核心，代表秩序的秩，描述的是两个随机变量的变化趋势方向和强度之间的关联。Spearman 相关系数是秩相关系数的一种，度量了两者单调关系的强弱，即两者在变大或变小的趋势上是否保持一致，在计算时使用的是数据的位次值。因此，Spearman 秩相关系数能用来表述非线性依赖关系，且不易受到异常值的干扰，具体计算公式如下：

$$\rho = \frac{\frac{1}{n} \sum_{i=1}^n (R(x_i) - \overline{R(x)}) \cdot (R(y_i) - \overline{R(y)})}{\sqrt{\left(\frac{1}{n} \sum_{i=1}^n (R(x_i) - \overline{R(x)})^2\right) \cdot \left(\frac{1}{n} \sum_{i=1}^n (R(y_i) - \overline{R(y)})^2\right)}} \quad (\text{公式 9})$$

其中， $R(x)$ 和 $R(y)$ 分别代表了 x 和 y 的位次， $\overline{R(x)}$ 和 $\overline{R(y)}$ 分别代表了两个变量的平均位次。

3. Kendall 秩相关系数

Kendall 相关系数是另一种秩相关系数，依据样本数据对之间的关系来进行判断，因此数据对也可以分为两类：一致对和分歧对。举例说明：变量 X 中连续的两个样本值记为 x_1 和 x_2 ，变量 Y 中连续的两个样本值记为 y_1 和 y_2 ，若两个变量这一对样本值取值的相对关系一致，即 $(x_2 - x_1)(y_2 - y_1) \geq 0$ 时，认为这一对样本值为一致对；若这一对样本值取值的相对关系不一致，即 $(x_2 - x_1)(y_2 -$

$y_1) \leq 0$ 时, 认为这一对样本值为分歧对。Kendall 相关系数的计算公式如下:

$$\tau_b = \frac{c - d}{\sqrt{(c + d + t_x)(c + d + t_y)}} \quad (\text{公式 10})$$

其中, c 和 d 分别代表了一致对和分歧对的个数, t_x 和 t_y 分别代表了变量 X 和变量 Y 中并列排位个数, 如果并列排位同时发生在两个变量中, 则即不计入 t_x , 也不计入 t_y 。

(二) 品种筛选

配对交易涉及双边买卖, 在挑选配对时, 为尽可能减少交易成本, 会选择流动性偏优的品种进行配对。因此, 首先依据流动性对品种进行筛选, 商品期货和金融期货分开考虑。同时, 为回测区间长度考虑, 只挑选早于 2016 年 1 月 1 日的品种进行配对。对商品流动性良好的品种定义为: 过去一年日均成交量不低于 6w 手且过去一个月日均成交量(双边)不低于 3 万手的品种。金融期货品种较少, 股指期货只有 4 种, 并且 IM 上市时间较晚, 因此仅在 IC、IF 和 IH 三个品种中进行选择, 国债期货也只有 4 种, TL 和 TS 上市时间较晚, 不进行配对考虑, 因此配对只有一对(T—TF)。最终筛选出的配对品种库如图表 2 所示。

图表 2: 配对品种库

商品期货					
板块	名称	上市时间	板块	名称	上市时间
农产品	豆油 Y	2006 年 1 月 9 日	黑色	螺纹钢 RB	2009 年 3 月 27 日
	菜油 OI	2007 年 6 月 8 日		热卷 HC	2014 年 3 月 21 日
	棕榈油 P	2007 年 10 月 29 日		硅铁 SF	2014 年 8 月 8 日
	豆粕 M	2000 年 7 月 17 日		锰硅 SM	2014 年 8 月 8 日
	菜粕 RM	2012 年 12 月 28 日		铁矿石 I	2013 年 10 月 18 日
	豆一 A	2002 年 3 月 15 日	能源	焦炭 J	2011 年 4 月 15 日
	玉米 C	2004 年 9 月 22 日		沥青 BU	2013 年 10 月 9 日
	玉米淀粉 CS	2014 年 12 月 19 日		PVC V	2009 年 5 月 25 日
	白糖 SR	2006 年 1 月 6 日		聚丙烯 PP	2014 年 2 月 28 日
	棉花 CF	2004 年 6 月 1 日		聚乙烯 L	2007 年 7 月 31 日
	鸡蛋 JD	2013 年 11 月 8 日		甲醇 MA	2011 年 10 月 28 日
有色	沪铜 CU	1993 年 3 月 31 日		橡胶 RU	1993 年 11 月
	沪铝 AL	1992 年 5 月 28 日		玻璃 FG	2012 年 12 月 3 日
	沪铅 PB	2011 年 3 月 24 日		PTA TA	2006 年 12 月 18 日
	沪锌 ZN	2007 年 3 月 26 日		焦煤 JM	2013 年 3 月 22 日
	沪锡 SN	2015 年 3 月 27 日	贵金属	沪银 AG	2012 年 5 月 10 日
	沪镍 NI	2015 年 3 月 27 日		沪金 AU	2008 年 1 月 9 日

金融期货			
板块	名称	上市时间	
股指	沪深 300IF	2010 年 4 月 16 日	
	中证 500IC	2015 年 4 月 16 日	
	中证 50IH	2015 年 4 月 16 日	

资料来源：中信期货研究所

（三）配对挑选规则

本次测试中所有配对采用**滚动式**判断，每 k 个所交易的配对依据前 1 个月的**日度收盘价**进行判断，商品期货每三个月进行一次轮动（每月 15 日，若 15 日不为交易日，提前至离 15 日最近的交易日），股指期货每一个月进行一次轮动。商品期货的具体配对选择步骤为：1. 对该板块所有品种价格取对数；2. 将任意品种两两配对算出各配对之间的 Pearson 相关系数、Spearman 秩相关系数和 Kendall 秩相关系数并进行相应排序；3. 将三种相关系数的排序值相加，取排名前 n 的配对进行交易。股指期货和商品期货一样的选择方法，但只取排名**第一**的配对进行交易。此外，在协整法中需对挑选出的配对进行平稳性检验，若无法通过检验，取排序后一位的配对。

五、策略构建和实证分析

本篇报告将对三类策略进行测试分析，分别为：协整策略、单一 Copula 函数策略和混合 Copula 函数策略。其中，协整策略最为简单和经典，仅需按照价格对数算出相应的买卖点就能进行交易，因此将协整策略作为对标策略，为后续优化提供参考。而在单一 Copula 函数策略中，需先对 Copula 函数进行最优选择再进行具体买卖点判断。最后，在混合 Copula 函数策略中，需先对混合权重进行推算再进行买卖点判断。

（一）回测说明

本篇报告中所有涉及回测的品种回测区间均为 2016 年 1 月 1 日至 2022 年 12 月 31 日，商品期货分板块测试，股指期货单独测试。商品期货分板块测试的原因是若将商品期货全部混合再根据相关性进行配对挑选，会造成**某一时期交易的品种集中于单一板块**，不利于分散风险和控制回撤。商品期货和股指期货的保证金比率统一设为 15%，保证金使用比率统一设为最高为 80%。回测总共使用 1000w 本金，各板块本金分配比例、手续费、挑选出的配对数和轮动周期如图表 3 所示，各板块的配对采用等额分配制度。此外，在农产品板块测试品种

相关性后发现，JD 和 SR 虽然在数理统计上与板块中的其他品种具有较强相关性，但关系极不稳定，因此在挑选配对时剔除了这两个品种。

图表 3: 回测说明

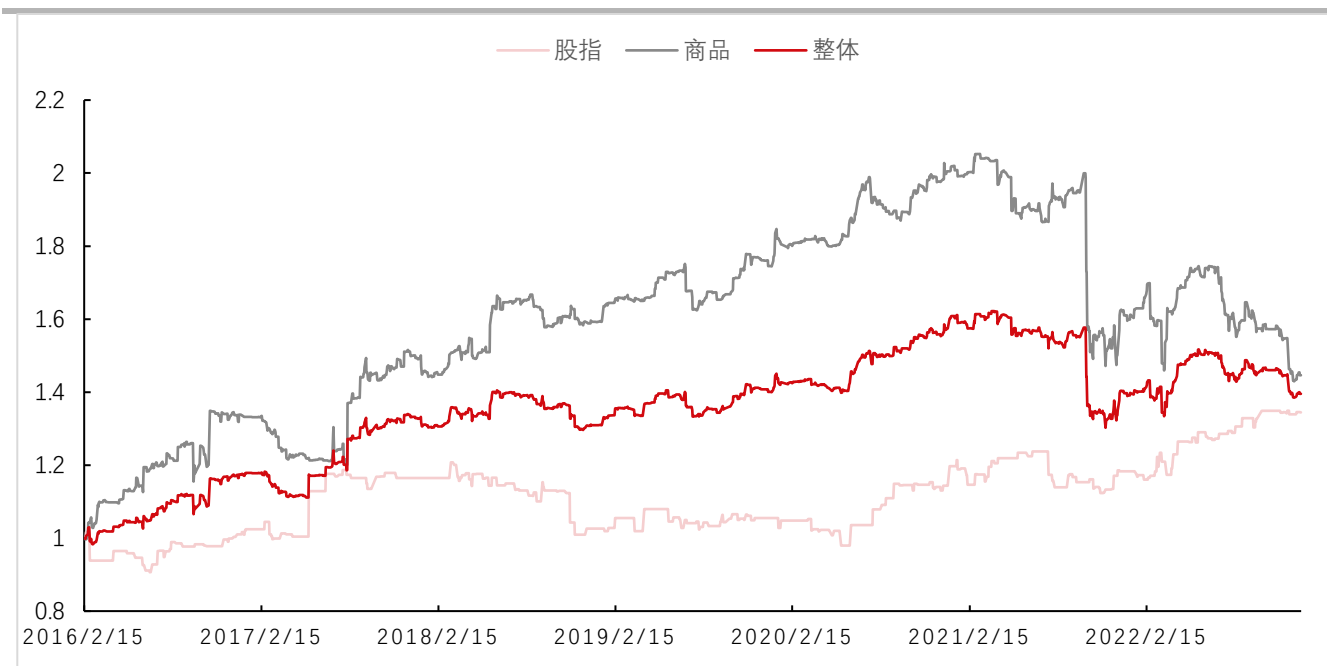
	农产品	有色	黑色	化工	贵金属	股指
本金	100w	150w	100w	100w	50w	500w
手续费	0.0003	0.0003	0.0003	0.0003	0.0003	0.0001
配对数	3	2	2	3	1	1
轮动周期	3 个月	3 个月	3 个月	3 个月	6 个月	1 个月

资料来源：同花顺 iFind、中信期货研究所

（二）协整策略

协整策略中的协整公式并非一成不变，因此采用滚动机制来不断调整协整配对系数和残差项，滚动周期与配对轮动周期一致，具体进出场点的判断依照上文所述，策略回测结果如图表 4 和图表 5 所示。

图表 4: 协整策略回测净值变化



资料来源：同花顺 iFind、中信期货研究所

图表 5: 协整策略回测结果

	总收益率	年化收益率	卡玛比率	最大回撤%	夏普比率
农产品	-25.36%	-3.72%	-0.17	49.65%	-0.17
有色	109.76%	16.11%	0.86	47.44%	0.47
黑色	17.75%	2.61%	0.10	27.41%	0.12
化工	56.73%	8.33%	0.11	77.81%	0.08

贵金属	18.60%	2.84%	0.04	70.87%	0.06
股指期货	34.47%	4.93%	0.26	18.90%	0.40
商品整体	44.61%	6.38%	0.21	30.34%	0.27
策略整体	39.54%	5.66%	0.29	19.70%	0.42

资料来源：同花顺 iFind、中信期货研究所

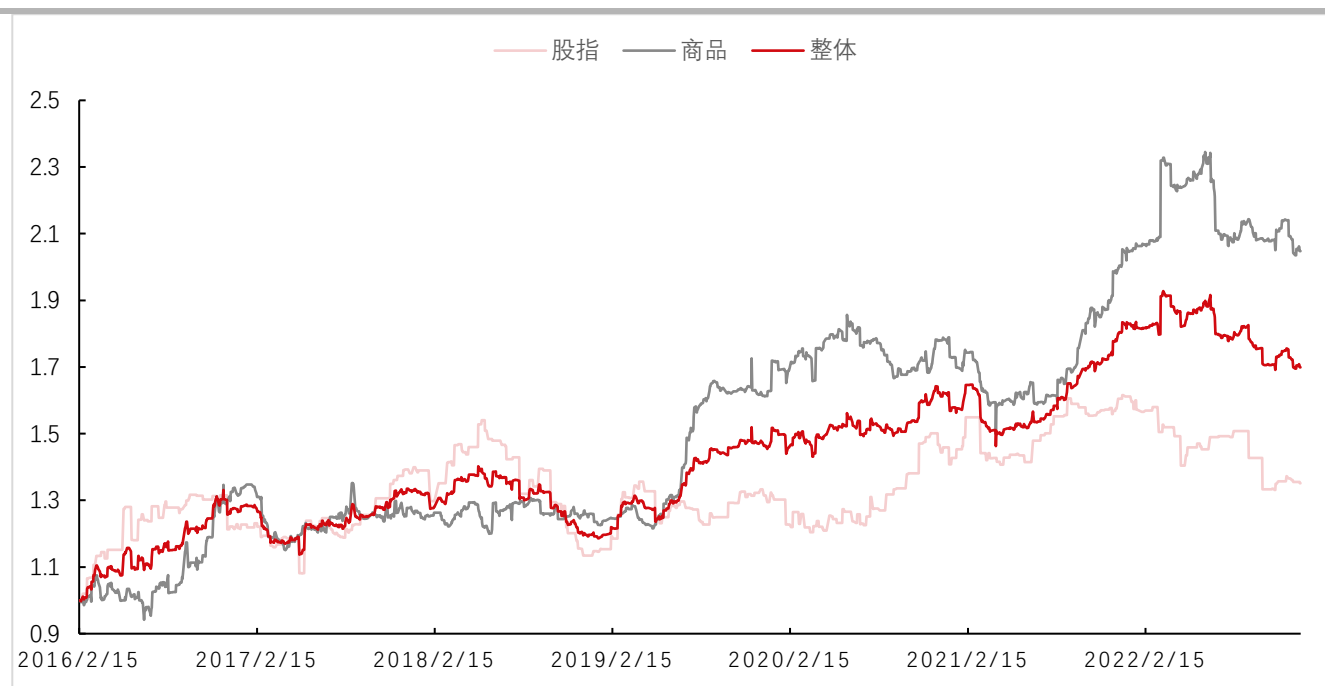
通过回测发现，协整策略在各个板块上运行结果差别较大，有色板块回测下来净值增长最多，农产品板块的净值则一致在水下，商品整体板块通过多板块配对组合控制了一定的最大回撤比例但仍然达到 30%，因此夏普和卡玛比率分别只有 0.27 和 0.21。股指期货部分由于最大回撤控制的比商品部分好，夏普和卡玛比率都高于商品板块。最终整体策略回测下来年化收益率达到 5.66%，夏普和卡玛比率分别为 0.42 和 0.29。协整策略的缺点很明显，**回测难以控制，交易轮次较少，交易信号对行情变化不够敏感**，不能及时对策略进行调整。

（三）单一 Copula 函数策略

单一 Copula 函数策略中 Copula 函数的选择采用了滚动机制，即凭借上一个区间内各函数的对数似然值 (loglike) 选择在下一个区间中使用 Clayton Copula、Gumbel Copula、Frank Copula 和 Gaussian Copula 函数中的任一函数，滚动周期与配对轮换周期一致。对数似然值可以用来判断 Copula 函数的拟合效果，值越大，拟合效果越好。本篇报告中 Copula 的拟合过程和参数计算全部采用的是 Copulae 库的相关函数，参数计算采用极大似然法。需要注意的是将序列投入 Copula 函数时需将序列转换成满足 (0, 1) 的均匀分布。具体进出场依据 Copula 函数的导数值 (Sign) 进行判断，当两个品种的收益率变化一致时，Sign 等于 0.5，所以当 Sign 大于 0.5 时做空配对；当 Sign 小于 0.5 时做多配对，当 Sign 回到 0.5 时平仓，公式推导如下：

$$\begin{aligned}
 \text{Sign} &= \frac{\partial C(u, v)}{\partial u} = p(R_t^X < r_t | R_t^Y = r_t) \\
 &= p(R_t^X - R_t^Y < r_t - r_t^Y | R_t^Y = r_t) \\
 &= p(\text{diff}_t^X < 0 | R_t^Y = r_t)
 \end{aligned}$$

图表 6：单一 Copula 函数策略回测净值变化



资料来源：同花顺 iFind、中信期货研究所

图表 7：单一 Copula 函数策略回测结果

	总收益率	年化收益率	卡玛比率	最大回撤%	夏普比率
农产品	8.28%	1.18%	0.03	46.29%	0.03
有色	124.48%	18.27%	0.73	24.98%	0.34
黑色	133.59%	19.61%	0.71	27.60%	0.53
化工	152.45%	22.38%	0.39	57.53%	0.38
贵金属	85.24%	12.51%	0.23	55.42%	0.35
股指期货	35.15%	5.03%	0.19	26.40%	0.25
商品整体	104.73%	14.98%	0.78	19.31%	0.66
策略整体	69.94%	10.00%	0.81	12.28%	0.65

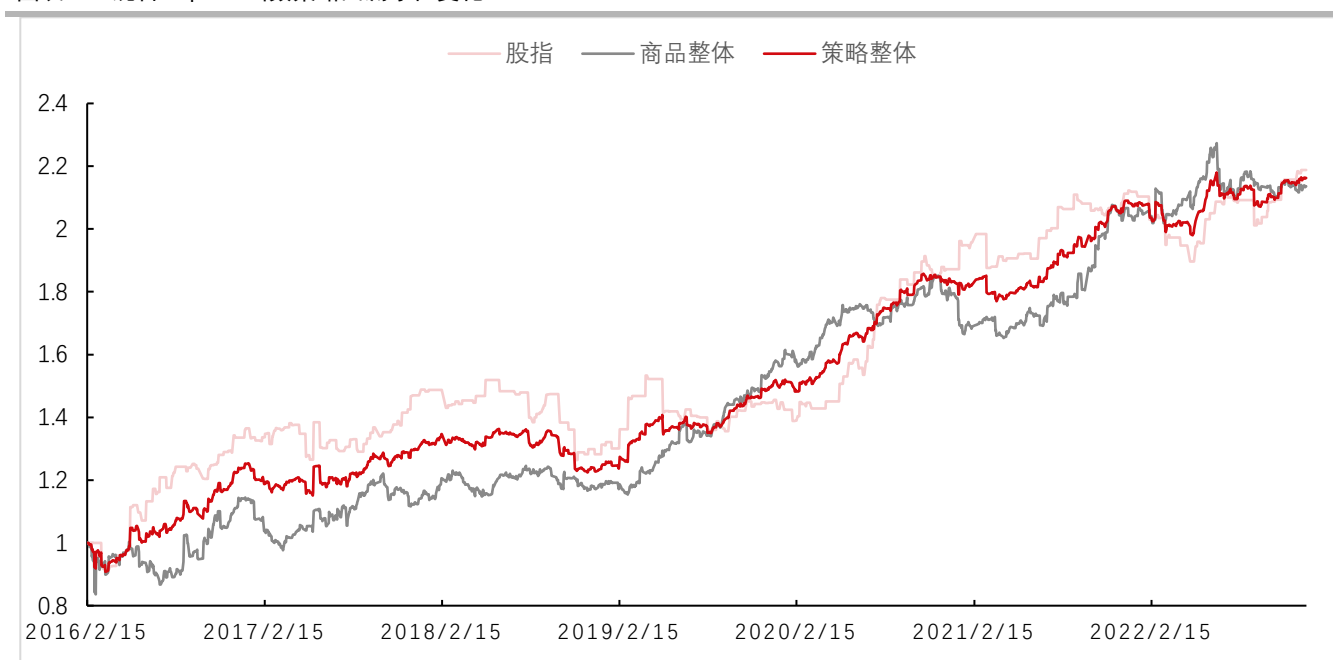
资料来源：同花顺 iFind、中信期货研究所

回测后发现，使用 Copula 进行入场判断不管在哪个板块进行回测，回测结果都比协整策略优秀，提高最为明显的就是商品板块，主要优化的点在于对最大回撤的控制，最终策略整体年化收益率达到 10%，夏普比率和卡玛比率分别为 0.65 和 0.81。Copula 从配对相关性变化进行入场考虑明显比简单利用配对量价关系进行入场判断灵敏许多，最大回撤、盈亏比和交易次数都提高很多。但单一的 Copula 函数在面对配对价差分布较为复杂的情况时，信号胜率会立刻下调，例如：在 2022 年下半年，商品期货趋势行情难出，频繁震荡的情况下，策略一直在回撤，因此单一 Copula 函数策略还有进一步优化的空间。

（四）混合 Copula 函数策略

混合 Copula 函数策略通过不同 Copula 函数的线性组合来增强分布的拟合效果和普适性，同样采用滚动机制，即在下一个区间使用上一个区间得出的混合 Copula 函数的线性拟合参数，滚动周期和配对轮换周期一致。指标计算方式和单一 Copula 函数策略一致，具体回测结果见图表 8 和图表 9。

图表 8：混合 Copula 函数策略回测净值变化



资料来源：同花顺 iFind、中信期货研究所

图表 9：混合 Copula 函数策略回测结果

	总收益率	年化收益率	卡玛比率	最大回撤%	夏普比率
农产品	97.31%	14.28%	0.81	17.73%	0.66
有色	37.52%	5.51%	0.10	55.81%	0.15
黑色	112.06%	17.71%	0.32	55.79%	0.34
化工	275.46%	40.43%	0.79	41.09%	0.98
贵金属	36.96%	5.64%	0.19	29.47%	0.21
股指期货	118.79%	16.99%	1.01	16.78%	0.86
商品整体	113.64%	16.25%	1.51	10.74%	0.86
整体策略	116.21%	16.62%	1.72	9.67%	1.22

资料来源：同花顺 iFind、中信期货研究所

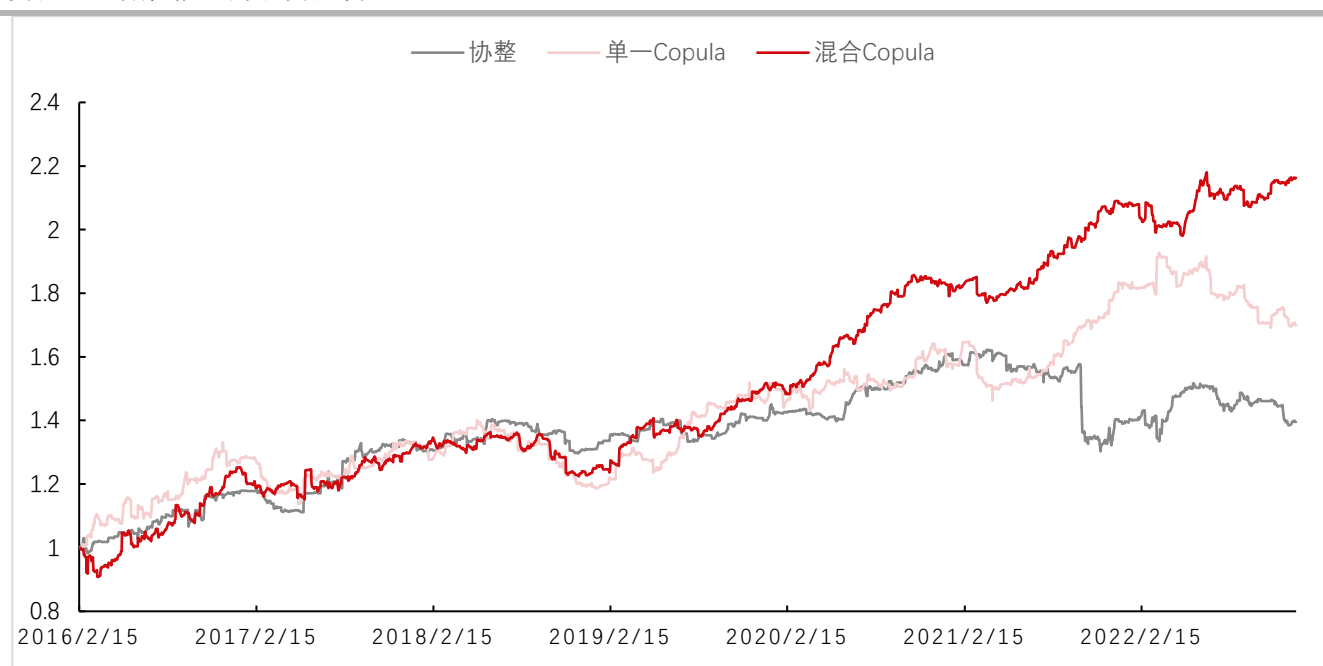
回测后发现，除了有色和贵金属板块，其他板块的回测结果都比较优异，商品整体表现和股指期货部分持平，策略整体年化收益率达到 16.62%，夏普和卡玛分别达到 1.22 和 1.72。通过将具有不同分布特点的 Copula 函数进行组合

增强了 Copula 函数的普适性，在面对变化较大的行情时，指标表现出了更强的敏锐性，因此收益率和最大回撤比率都得到了优化。

六、结论与展望

本文首先详细介绍了经典协整策略和相关的 Copula 函数要点，然后针对 Copula 函数提出了两种创新型配对策略：单一 Copula 策略和混合 Copula 策略，并在商品期货和股指期货上进行了相应的回测。在进行配对选择时，本文使用了轮动机制来提高配对库的时变性；同时，计算了 Pearson、Spearman 和 Kendall 三大相关系数尝试从不同角度刻画配对的相关性。通过对三个策略进行回测，我们发现以下几点：

图表 10：各策略回测净值变化综合



资料来源：同花顺 iFind、中信期货研究所

图表 11：回测结果总结

	总收益率	年化收益率	卡玛比率	最大回撤%	夏普比率
协整	39.54%	5.66%	0.29	19.70%	0.42
单一 Copula	69.94%	10.00%	0.81	12.28%	0.65
混合 Copula	116.21%	16.62%	1.72	9.67%	1.22

资料来源：同花顺 iFind、中信期货研究所

1. 使用 Copula 函数对配对进行进出场判断效果明显优于协整策略，混合 Copula 函数策略的年化收益率达到 16.62%，较协整策略提高了 10% 以上；夏普

比率和卡玛比率分别达到 1.22 和 1.72，是协整策略的 2.9 倍和 5.9 倍。

2. 混合 Copula 函数策略融合了三大阿基米德 Copula 函数的特点，在面对复杂的现实行情时，较单一 Copula 函数体现出更强的灵活度和敏锐性。

3. 不同板块的配对回测情况差别较大，通过多板块组合可以进一步控制回撤程度，提高整体策略的稳定性。同时，并不是在所有板块中，混合 Copula 函数回测结果都显著高于其他策略，但混合 Copula 函数拥有更强的普适性。

4. 在三个策略中，化工板块都表现出较强的盈利能力，因此推测化工的板块品种具有更强的联动性，在进行配对交易时，可以尝试增强在此板块的投入来优化策略的整体净值。

虽然此次回测的效果较好，但本次测试也自带局限性。首先，在本金分配比列和配对数量选择上依靠作者主观意念进行考虑，并未采用量化风控手段。其次，混合 Copula 策略的组合方式为简单的线性拟合，并未考虑其他多维的组合手段。第一点在后续报告中会有进一步的优化导论，第二点可以作为该策略进一步优化重点。

参考文献

- 1) STANDAER Y, MARAIS D, BOTHA I. Trading strategy with Copulas[J]. Journal of Economics and Financial Sciences, 6(1):83-108, 2013.
- 2) 张戈, 程棵, 陆凤彬等. 基于 Copula 函数的程序化交易策略[J]. 系统工程理论与实践, 2011, 31(4): 599-605.
- 3) 沈银芳, 郑学东, 徐信喆. 基于混合 Copula 的 ETF 配对交易策略[J]. 浙江大学学报. 2016, 43(3).
- 4) 沈银芳, 郑学东, 徐建军. 基于时变混合 Copula 模型的配对交易策略[J]. 财经论丛. 2016, 212(10).
- 5) 仇中群, 程希骏. 基于协整的股指期货跨期套利策略模型[J]. 财系统工程. 2008, 26(12).
- 6) 张俊, 李妍. 配对交易策略: 一个文献评述[J]. 中国证券期货. 2010, pp: 28-29.
- 7) 张青, 基于 Copula 理论的配对交易策略. [基于 Copula 理论的配对交易策略 - 中国知网 \(cnki.net\)](#).
- 8) 何卫平. 基于 Copula 函数和条件概率模型的配对交易研究——以我国沪深 300 指数成分股为例. [基于 Copula 理论的配对交易策略 - 中国知网 \(cnki.net\)](#)

免责声明

除非另有说明，中信期货有限公司拥有本报告的版权和/或其他相关知识产权。未经中信期货有限公司事先书面许可，任何单位或个人不得以任何方式复制、转载、引用、刊登、发表、发行、修改、翻译此报告的全部或部分材料、内容。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中信期货有限公司所有或经合法授权被许可使用的商标、服务标记及标记。未经中信期货有限公司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。

如果在任何国家或地区管辖范围内，本报告内容或其适用与任何政府机构、监管机构、自律组织或者清算机构的法律、规则或规定内容相抵触，或者中信期货有限公司未被授权在当地提供这种信息或服务，那么本报告的内容并不意图提供给这些地区的个人或组织，任何个人或组织也不得在当地查看或使用本报告。本报告所载的内容并非适用于所有国家或地区或者适用于所有人。

此报告所载的全部内容仅作参考之用。此报告的内容不构成对任何人的投资建议，且中信期货有限公司不会因接收人收到此报告而视其为客户。

尽管本报告中所包含的信息是我们于发布之时从我们认为可靠的渠道获得，但中信期货有限公司对于本报告所载的信息、观点以及数据的准确性、可靠性、时效性以及完整性不作任何明确或隐含的保证。因此任何人不得对本报告所载的信息、观点以及数据的准确性、可靠性、时效性及完整性产生任何依赖，且中信期货有限公司不对因使用此报告及所载材料而造成的损失承担任何责任。本报告不应取代个人的独立判断。本报告仅反映编写人的不同设想、见解及分析方法。本报告所载的观点并不代表中信期货有限公司或任何其附属或联营公司的立场。

此报告中所指的投资及服务可能不适合阁下。我们建议阁下如有任何疑问应咨询独立投资顾问。此报告不构成任何投资、法律、会计或税务建议，且不担保任何投资及策略适合阁下。此报告并不构成中信期货有限公司给予阁下的任何私人咨询建议。

深圳总部

地址：深圳市福田区中心三路8号卓越时代广场（二期）北座13层1301-1305、14层

邮编：518048

电话：400-990-8826

传真：(0755) 83241191

网址：<http://www.citicsf.com>