



北京大学
PEKING UNIVERSITY

基于 Copula 函数的配对交易研究

唐熙勳 2301210041

2023 年 12 月 19 日



在金融市场中，配对交易是一种常见的交易策略，旨在通过同时买入和卖出两个或多个相关性较高的金融资产，从中获得收益。传统的配对交易方法通常基于统计模型或技术指标，如协整关系或均值回归，来识别资产价格之间的关联性。

然而，传统的配对交易方法在面对非线性关系或复杂市场条件时可能存在一定的局限性。为了克服这些限制，Copula 函数作为一种灵活的工具被引入到配对交易中。Copula 函数是一种用于描述多维随机变量之间依赖关系的数学函数，它能够捕捉到非线性和非正态的相关性结构。



Copula 函数的一个性质

记 X, Y 的分布函数分别为 $F_1(x), F_2(x)$, 令
 $U = F_1(X), V = F_2(Y)$, 联合分布函数记为 F , 根据 Copula 函数
定义, $P(X \leq x, Y \leq y) = F(x, y) = c(F_1(x), F_2(y))$
从而条件分布

$$\begin{aligned} P(X \leq x | Y = y) &= \frac{P(X \leq x, Y = y)}{P(Y = y)} = \frac{P(U \leq u, V = v)}{P(V = v)} \\ &\approx \frac{P(U \leq u, V \in (v + \Delta v))}{P(V \in (v + \Delta v))} \\ &= \frac{c(u, v + \Delta v) - c(u, v)}{\Delta v} \\ &\rightarrow \frac{\partial c}{\partial v}(u, v) \quad (\Delta v \rightarrow 0) \end{aligned}$$



类似地,

$$P(Y \leq y | X = x) \approx \frac{\partial c}{\partial u}(u, v)$$

$$P(X \leq x | Y = y) \approx \frac{\partial c}{\partial v}(u, v)$$

在估计出 Copula 参数后, 可以通过取 $\Delta u, \Delta v$ 较小的值来近似条件概率。



1. 寻找股票对：基于 Pearson、Kendall、Spearman 等相关系数。
2. 拟合 Copula 函数：假设边缘分布为 t 分布，联合 Copula 为 t -Copula，使用 IFM 方法估计参数。
3. 产生交易信号：通过计算条件概率判断股票对之间的相对高估与低估，并以此产生交易信号。

数据方面，使用 2014 年 1 月沪深 300 成分股在 2014 年 1 月到 2018 年 1 月的数据，同时，以月为单位滚动更新对股票对的二元 Copula 函数估计，作为生成下一个月中每日交易信号的依据。



在给定的股票池与区间，计算 Pearson 相关系数矩阵、Kendall's τ 相关系数矩阵、Spearman's ρ 相关系数矩阵，找出各矩阵中相关系数大的部分的股票对，在三者的并集中筛选。

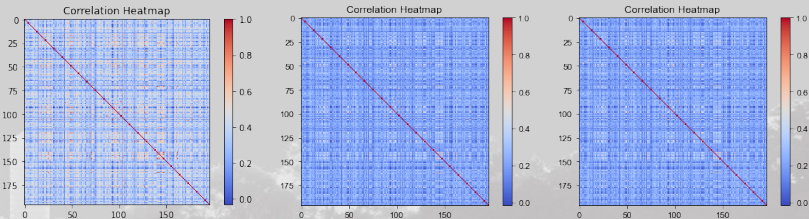


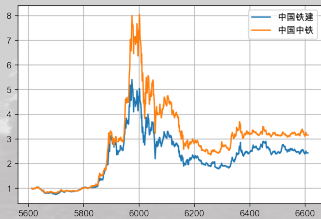
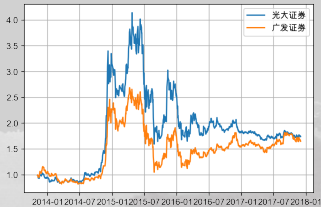
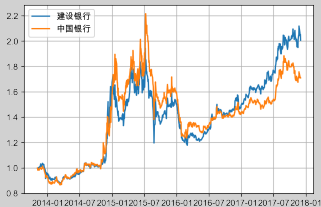
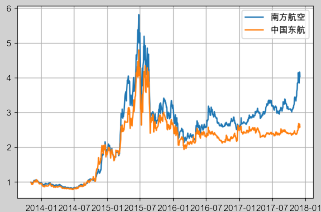
图: 从左到右 Pearson 相关系数矩阵、Kendall's τ 相关系数矩阵、Spearman's ρ 相关系数矩阵



在给定的股票池与区间，计算 Pearson 相关系数矩阵、Kendall's τ 相关系数矩阵、Spearman's ρ 相关系数矩阵，找出各矩阵中相关系数大的部分的股票对，在三者的并集中筛选。

表: 股票对相关系数

1	2	Pearson	Kendall	Spearman
南方航空	中国东航	0.88	0.67	0.84
建设银行	中国银行	0.85	0.66	0.82
光大证券	广发证券	0.89	0.67	0.82
中国铁建	中国中铁	0.89	0.71	0.86



图：股票对的价格走势（归一）



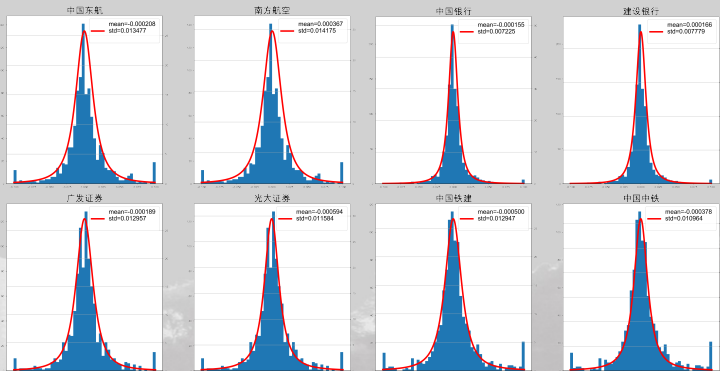
本文采用 IFM 方法，首先根据股票对历史收益率数据拟合收益率的分布，再估计 Copula 函数的参数。

具体地说，根据股票收益率“尖峰厚尾”的特征，假设收益率服从 t 分布，采用 IFM 方法，使用最大似然估计方法估计分布函数的参数；得到边缘分布函数后，将股票对收益率序列转换为 $[0, 1]$ 上的均匀分布序列，随后再次使用最大似然估计方法估计 t-Copula 函数的参数。

$$C(u, v, \rho) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{r^2 + s^2 - 2\rho rs}{2(1-\rho^2)}\right] dr ds$$



拟合 Copula 函数



图：股票对的收益率拟合



表: 估计 t-Copula 参数

	股票 1	股票 2	ρ	ν
1	南方航空	中国东航	0.692	14.71
2	建设银行	中国银行	0.741	8.28
3	光大证券	广发证券	0.740	5.76
4	中国铁建	中国中铁	0.753	9.71



在确定股票对收益率的边缘分布与 Copula 函数的估计后，便可以根据新数据的输入得到交易信号，具体地，考虑如下信号

$$\begin{aligned} \text{signal}_1(u_1 | u_2) &= P(U_1 \leq u_1 | U_2 = u_2) = \frac{\partial C(u_1, u_2)}{\partial u_2}, \\ \text{signal}_2(u_2 | u_1) &= P(U_2 \leq u_2 | U_1 = u_1) = \frac{\partial C(u_1, u_2)}{\partial u_1}. \end{aligned}$$

其中， u_1, u_2 为新一日收益率数据 r_1, r_2 分别通过对应的边缘分布作用得到。偏导数通过取较小的 $\Delta u, \Delta v$ 值近似。



$$\begin{aligned} \text{signal}_1(u_1 | u_2) &= P(U_1 \leq u_1 | U_2 = u_2) = \frac{\partial C(u_1, u_2)}{\partial u_2}, \\ \text{signal}_2(u_2 | u_1) &= P(U_2 \leq u_2 | U_1 = u_1) = \frac{\partial C(u_1, u_2)}{\partial u_1}. \end{aligned}$$

signal 值在 $[0, 1]$ 之间，当 signal 值趋于 1 时，说明通过 Copula 函数求偏导给出对应的条件概率偏大，前者收益率高估，前者股票较后者股票高估；当 signal 值趋于 0 时，说明通过 Copula 函数求偏导给出对应的条件概率偏小，前者收益率低估，前者股票较后者股票低估。

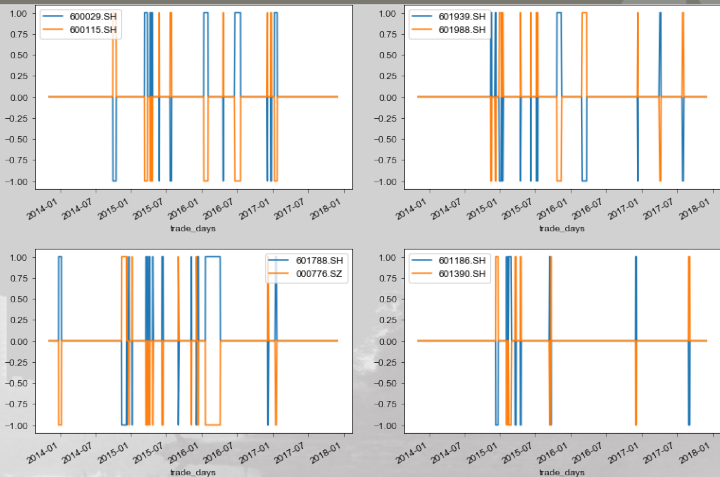


本文中，交易信号如下生成：

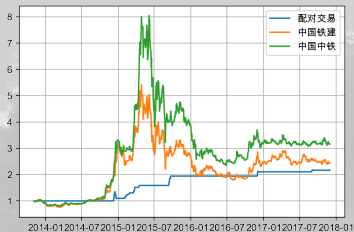
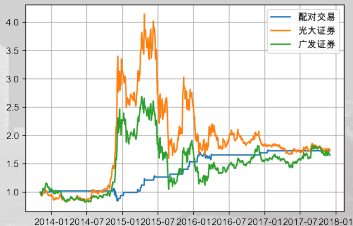
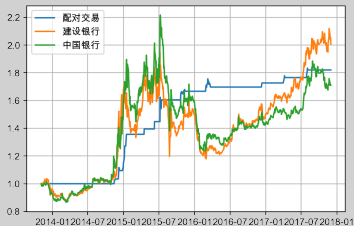
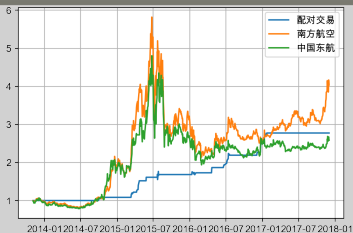
- ▶ 当 signal1 值大于 0.9 且 signal2 值小于 0.1 时，做多股票 1，做空股票 2；
- ▶ 当 signal1 值小于 0.6 且 signal2 值大于 0.4 时，如果处于做多股票 1，做空股票 2 的状态，则平仓；
- ▶ 当 signal1 值小于 0.1 且 signal2 值大于 0.9 时，做多股票 2，做空股票 1；
- ▶ 当 signal1 值大于 0.4 且 signal2 值小于 0.6 时，如果处于做多股票 2，做空股票 1 的状态，则平仓；



生成交易信号



图：配对交易信号



图：配对交易回测结果



表：配对交易回测结果（年化）

	1	2	收益率	标准差	夏普比
0	南方航空	中国东航	0.4439	0.2698	1.65
1	建设银行	中国银行	0.2053	0.1934	1.06
2	光大证券	广发证券	0.1827	0.2762	0.66
3	中国铁建	中国中铁	0.2919	0.2850	1.02

可以发现，配对交易方式虽然在总体收益上可能低于最简单的买入并持有策略，但其清晰的交易信号控制了回撤的范围，在大部分时间都处于空仓等待交易机会的阶段，在股票对出现相对的高估低估后开仓，胜率较高，在夏普比上便体现为较高的值。



Copula 理论能很好地应用于金融市场间的相关性分析、投资组合分析、风险管理、波动溢出分析、资产定价、市场风险分析、信用风险分析。

本文将 Copula 理论应用于配对交易中，在 A 股市场沪深 300 股票池中寻找配对交易的股票对，并基于 Copula 函数进行配对交易，实现了近似套利交易的低回撤高夏普的收益，说明 Copula 理论的普遍性与实用性。



北京大学
PEKING UNIVERSITY

谢谢!