

基于GARCH模型的商品期货跨期套利研究

根据置信水平调整阈值界限,以满足不同风险偏好投资者的需求

■ 闫丰 姚禹

对期货市场来说,跨期套利能够增加期货合约的流动性,促使合约间价差在合理区间内运行。棉花作为关系国计民生的重点品种,其价格波动程度关系到整个棉纺产业链的平稳发展。因此,本文以棉花期货为标的,通过GARCH套利模型,建立适合不同风险偏好投资者的跨期套利策略,以期丰富相关套利研究成果。

研究背景

跨期套利,是指同一期货品种不同合约之间,利用其价差变化,建立数量相等、头寸方向相反,最终以对冲或者交割形式了结的一种套利策略。不同于一般的投机交易,跨期套利交易因其同时买卖方向相反的同种合约,具有对冲特性,风险较低。同时,由于套利交易所构建的是一种价差组合,相对于单边投机,其价差波动率幅度更小,整体收益的稳定性更高。除此以外,对期货市场来说,跨期套利活动的存在能够增加期货合约的流动性,促使合约间价差在合理区间内运行,保障期货市场发现价格的功能。因此,本文从理论层面对跨期套利进行研究,具有一定的意义。

从模型层面来说,金融时间序列具有异方差性特征。就跨期套利而言,其通常表现为价差具有波动集聚的特点,即价差高的现象集聚在一起,而价差低的现象集聚在一起。针对异方差性,采用GARCH模型建模是最合适的。该模型主要认为,当某一组合的价差偏离标准达到一定程度时,即会出现套利机会,此时便可建仓入场,等到价差回归到均衡时平仓做反方向交易获利。因此,本文采用GARCH模型对套利交易进行研究。

本文选定棉花期货作为研究对象。棉花期货的交易制度、监管体系、合约活跃度已得到较为完善的发展,市场认可度较高。棉花作为关系我国民生的重点品种,其价格波动程度关系到整个棉纺产业链的平稳发展。自2021年棉价大涨之后,棉市呈现出反向特征,即近月合约价格大于远月合约价格。1月合约、5月合约与9月合约之间的价差也都脱离正常区间,出现反向套利空间,具体走势情况如图1和图2所示。

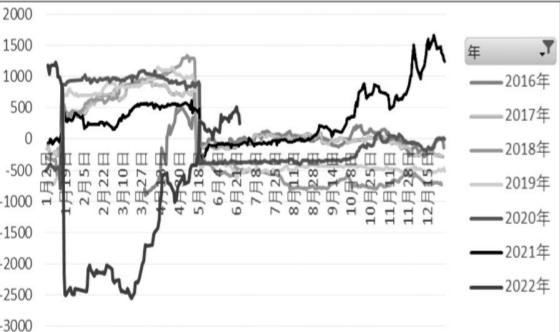


图1为CF01—05合约价差走势

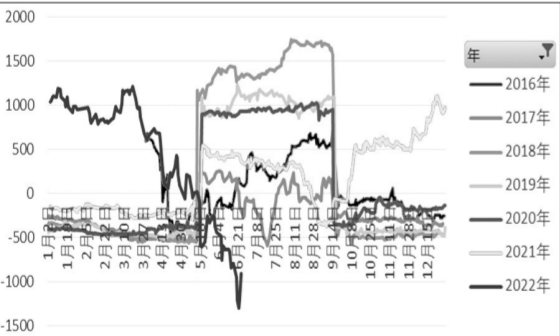


图2为CF05—09合约价差走势

随着反向跨期套利者逐步介入,反向套利空间逐步缩小,合约价差也逐步向合理区间回归,推动棉价回落。因此,本文以棉花期货为标的,通过GARCH套利模型建立适合不同风险偏好投资者的跨期套利策略,以期丰富相关套利研究成果。

理论分析

基于传统持有成本理论的跨期套利理论认为,商品期货在t时刻的价格可以用公式(1)表示:

$$F_t = S_t \times e^{r(T-t)} + C \quad (1)$$

其中,F表示在t时刻期货价格,S代表在t时刻的现货价格,T为期货合约最后交易日,r代表无风险利率。C代表持有成本,包括仓库费用、保险费用、储存损耗等。根据公式(1)计算出来的期货价格为理论均衡价格,但根据传统持有成本理论,我们无法解释期货价格呈现反向市场的原因。因此,需要引入“便利收益”的概念。期货市场参与者包括套期保值产业客户与普通投机者。套期保值者在期货市场上进行风险规避的操作时,需要付出相应的风险溢价R作为对投机者承担特定风险的补偿,以换取自身期现整体利润水平的稳定。当

套期保值者在期货市场上进行卖出操作时,此时风险溢价R显示负值;当套期保值者在期货市场上进行买入操作时,风险溢价R为正值。正是由于风险溢价R的存在,期货市场才拥有正向与反向之分。因此,我们对传统持有成本理论下的期货定价公式进行修正,得出公式(2):

$$F_t = S_t \times e^{r(T-t)} + C + R \quad (2)$$

对于同标的不同合约之间的期货价差,可由公式(3)表示:

$$\Delta F = S_t \times e^{r \times (T_2 - t)} - S_t \times e^{r \times (T_1 - t)} \quad (3)$$

其中,ΔF代表合约理论价差,T2表示远月合约最后交易日,T1表示近月合约最后交易日。若远月合约与近月合约实际价格之差在剔除相关成本之后,仍然大于合约理论价差,即当公式(4)成立,则表明存在套利空间:

$$F(S, t, T_2) - F(S, t, T_1) \geq \Delta F + C + R \quad (4)$$

F(S,t,T₁)表示在t时刻期货F的。此时可采取牛市套利策略,即买入近月合约卖出远月合约,当实际价差向理论价差回归时可平仓获利。

本文所建立的GARCH模型以修正的持有成本理论为基础,从属于统计套利理论,其本质上属于中性策略。不同于投机交易依靠对后市标的行情走势方向判断获利,统计套利并不依赖于标的绝对方向变化,其核心原则是基于标的合约不同月份之间的价差,具有均值回复性和相对平稳性的特征。

通过统计分析,在确定标的合约具有长期稳定的协整关系后,当价差序列出现短暂的偏离时,即可通过合约之间的反方向交易,等待价差序列回归正常区间时平仓,从而获取相关利润。从数学的角度来看,对统计套利的定义如下:

假设某个交易策略x_t的初始投入为0,该交易策略在未来t时刻经过无风险利率折现之后,其价值为v(t)。若v(t)符合下列三项条件,则证明符合统计套利的特征。

- 第一项条件是v(0)=0,即初始投资为0。
- 第二项条件是 $\lim_{t \rightarrow \infty} E[v(t)] > 0$,表示随着时间推移,期望利润回报大于0。
- 第三项条件是 $\lim_{t \rightarrow \infty} P[v(t) < 0] = 0$,表示随着时间推移,该交易策略发生亏损的可能性近似于0。

GARCH模型建立的基本原理是两个月份的股指期货合约价格之间的相关关系出现差异,预期未来价差将回归理性。根据协整理论,跨期套利成功的前提是两合约之间的价差存在协整关系。当观测到价差偏离一定程度时开始建仓,即买入被低估的资产,卖出被高估的资产。等到价差回归理性区间时,止盈离场。若价差进一步扩大到一定程度时,止损出局。

根据统计套利的协整策略,我们可按如下步骤构建套利模型,实时监控棉花期货跨期套利的机会,并在合适的时间实施套利交易策略:

- 第一步,对所选合约的历史价差分布进行相关分析。
- 第二步,对两个到期日不同、标的物相同的期货合约的价格序列进行平稳性检验。若原序列不平稳,而其一阶差分序列平稳,则转入下一步。
- 第三步,对两个价格序列进行协整关系检验。
- 第四步,对协整方程的残差序列进行ARCH-LM检验。若存在ARCH效应,可以采用ARCH模型刻画残差的异方差性。
- 第五步,序列存在ARCH效应的条件下,对两个价格序列运用GARCH模型拟合。
- 第六步,利用拟合后的GARCH模型,算样本期外的实际价差和对应的方差。
- 第七步,设定阈值,建立入场与出场临界点,依据套利结果评价套利模型。

实证检验

考虑到合约的流动性,本文以CF01合约与CF05合约价差为一组,计为C1;CF05合约与CF09合约价差为一组,计为C2;CF01合约与CF09合约价差为一组,计为C3。总共3组价差作为研究样本,所选取样本区间为2016年1月1日至2022年7月6日,总共1583个交易日的数据,建立跨期套利交易模型。

我们对3组时间序列进行相关性分析,初步判断3组数据是否具有协整关系。根据计算可得,C1组之间的相关系数为0.961755,C2组之间的相关系数为0.962022,C3组之间的相关系数为0.969719,基本上可近似看作完全正相关关系。因此,3组时间序列很大程度上具有协整关系。



为防止出现伪回归现象,需要对样本数据进行平稳性检验。为降低数据样本的波动性,我们将其进行对数化处理后进行ADF检验,检验结果如表1所示。

变量	ADF值	1%临界值	5%临界值	10%临界值
D(C1)	-33.66732	-3.434281	-2.863163	-2.567682
D(C2)	-33.47428	-3.434281	-2.863163	-2.567682
D(C3)	-33.51055	-3.434281	-2.863163	-2.567682

表1为ADF检验结果

表中D(C_i)代表对变量进行一阶差分。若ADF值大于临界值,则表明时间序列不平稳;若ADF值小于临界值,则证明该时间序列具有平稳性。由表1可知,3组样本原时间序列均不满足平稳性特征,但其一阶差分序列在1%临界值的显著水平下是平稳的,即一阶单整序列通过ADF检验,满足协整检验的要求。

接下来,通过Engle-Granger对样本进行协整检验,第一步对3组数据建立OLS估计方程,表示残差项,其结果如下所示:

$$C1: CF05 = 0.958414 \times CF01 + 0.401656 + \mu_{t1}$$

$$R\text{-squared} = 0.928782 \quad F\text{-statistic} = 20618.40$$

$$C2: CF09 = 0.9689 \times CF05 + 0.293071 + \mu_{t2}$$

$$R\text{-squared} = 0.936082 \quad F\text{-statistic} = 23153.72$$

$$C3: CF09 = 0.961982 \times CF01 + 0.361 + \mu_{t3}$$

$$R\text{-squared} = 0.933032 \quad F\text{-statistic} = 22027.31$$

因为3组数据OLS方程中的R-squared值接近于1,所以可以认为OLS的拟合效果较好。不过,通过观察估计结果的残差图(图3、图4、图5),可以发现,残差可能存在异方差。为进一步证明这种异方差性,我们需要对残差进行ARCH效应检验。

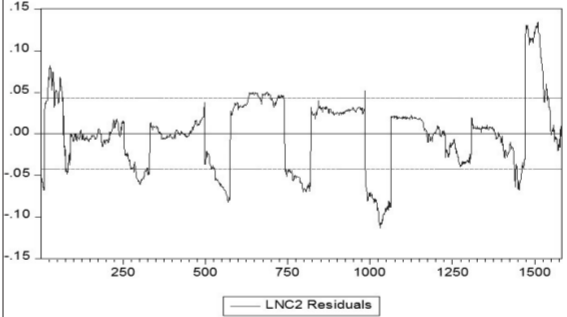


图3为C1残差

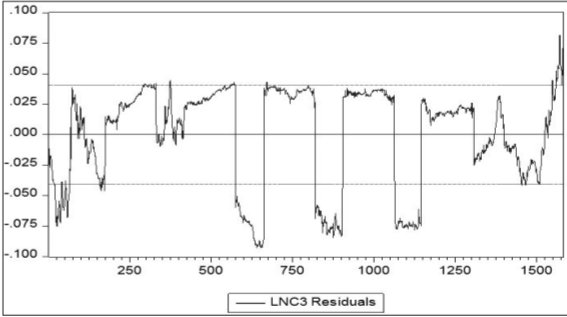


图4为C2残差

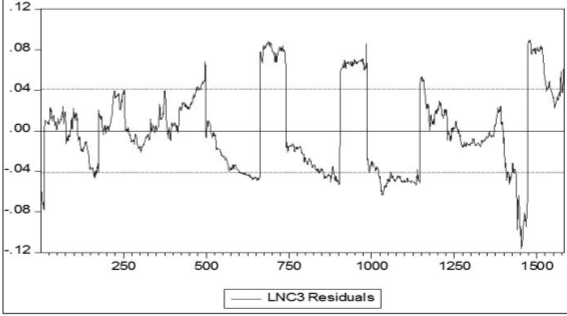


图5为C3残差

由表2可知,3组时间序列Obs*R-squared都处于较大位置,同时P值为0。表明3组时间序列估计的残差存在ARCH效应。因此,可以采用GARCH模型来描述样本序列。

ARCH Test	F-statistic	Obs*R-squared	Prob. F(1,1580)	Prob. Chi-Square(1)
C1	48957.27	1532.540	0.0000	0.0000
C2	38608.99	1519.805	0.0000	0.0000
C3	37407.42	1517.888	0.0000	0.0000

表2为残差ARCH检验结果

残差序列存在ARCH效应的前提下,我们使用GARCH(1,1)模型来刻画残差序列的条件异方差性,均值方程和条件方差方程中,所有参数估计值的统计量非常显著,其相应的概率值非常小,说明这些参数估计值都是显著的。均值方程中的参数估计值都大于0,从而保证条件方差的非负数要求符合模型参数要求。因此,可以写出3组样本的模型估计结果。

对于C1组样本数据,均值方程CF05=0.949533×

CF01+0.484004。

条件方差σ_t²=2.69E-06+0.701484u_{t-1}²+0.281493σ_{t-1}²。

对于C2组样本数据,均值方程CF09=0.962817×

CF05+0.369556。

条件方差σ_t²=7.00E-06+0.772057u_{t-1}²+0.010874σ_{t-1}²。

对于C3组样本数据,均值方程CF09=0.98049×

CF01+0.868116。

条件方差σ_t²=3.64E-06+0.720520u_{t-1}²+0.033757σ_{t-1}²。

记所估计的GARCH模型的残差序列为Resid(i)(i=1,2,3),条件方差序列为Garch(i)(i=1,2,3)经检验,序列是平稳序列,并且不存在ARCH效应,这也说明模型的拟合效果较好。接下来,我们构造一新残差序列RESID(i)= $\frac{\text{Resid}(i)}{\sqrt{\text{Garch}(i)}}$ (i=1,2,3),观察序列的统计特征,有关统计量的值见表3。

序列	均值	标准差	偏度	峰度	Jarque-Berate	概率
RESID(1)	-0.06871	0.998165	0.08168	2.681191	0.84751	0.39
RESID(2)	-0.05813	0.981619	0.08815	2.691605	0.83894	0.37
RESID(3)	-0.06818	0.996158	0.08384	2.671916	0.84891	0.39

表3为序列RESID相关统计量

由上表可知,3组序列均不能拒绝序列服从正态分布原假设,即序列RESID服从正态分布。再对序列RESID进行均值检验和方差检验,结果证明序列RESID服从均值为0、方差为1的原假设。由此可以认为,序列近似服从标准正态分布。

接下来,我们对数据进行回测检验。设建仓阈值为u₁,平仓阈值为u₂,止损阈值为u₃。为避免冲击成本的产生,套利时只考虑1手合约,建立如下跨期套利交易策略:当实际价差高于理论价差,即RESID>u₁时,买入1手近月合约,同时卖出1手远月合约。当RESID回归至(-u₂,u₂)时,平仓获利出场。若实际价差不收敛时,RESID>u₃,止损出局。

对于阈值u₁、u₂和u₃的选取,我们利用标准正态分布的置信水平来确定其取值,这样可以尽可能规避主观判断对选取价差的影响。置信水平的高低选取能反映投资主体对风险的厌恶程度。选择的置信水平越高,厌恶风险的程度越大。对于风险厌恶者,可选取高置信水平,进一步规避风险。

最后,我们利用样本内的数据对模型进行实证检验。为方便计算,选取u₂=0.25,u₃=2,只对u₁取置信水平。本文规定u₁的置信水平不低于90%,这样可以在90%的概率下保证套利交易在下一交易日不亏损。接下来,我们以90%的置信度(对应的标准正态分布数1.28)进行数据回测,同时考虑相关交易手续费,结果如下表所示:

	样本内	样本外
套利总次数	766	322
套利成功次数	652	265
套利成功率	85.12%	82.6%
净盈利	5941.64	3719.89
平均年化收益	6.87%	4.61%
最大回撤率	11%	13%

表4为样本回测结果

通过表4的回测结果可以看出,本文所建立的GARCH套利模型具有较好的拟合效果。整体套利成功率较高,最大回撤也处于可接受水平。就样本外数据检验而言,该模型也具有较高的契合度,但表现程度不如样本内数据,依然具有一定的改进空间。

结论建议

本文基于GARCH模型,对以棉花期货为标的的跨期套利模型进行了深入分析与实证检验。在进行套利回测检验时,通过置信水平来确定套利模型的阈值界限。置信水平越高,套利机会越少,但风险越小。即本文所建立的套利模型对不同的风险偏好投资者均适用。

本文以90%的置信度为衡量准绳,分样本内、样本外进行跨期套利交易。结果显示,棉花期货市场在样本内、样本外均存在较多跨期套利机会。除此之外,该模型套利成功率较高,能够获得较为可观的收益。

总体来看,基于GARCH模型的棉花期货跨期套利交易具有可落地性和收益性,但平均年化收益率对部分投资者来说不具有吸引力,需要对模型策略进一步改进。正是期货市场活动中套利者的存在,才推动期现价差、跨期价差在合理区间运行,避免长时间面临极端价差行情,保证期货市场更好地发挥其价格发现功能。

(作者单位:华安期货)