

股指期货跨期套利策略的比较

——基于统计套利视角

摘要:

- ◆ 本文比较两类不同的跨期套利方法及其收益效果：一类是基于持有成本模型；一类是基于统计套利。其中统计套利模型分历史价差法、协整模型和 GARCH 模型三种方法分别进行计算。
- ◆ 从价差的计算方法来看，持有成本模型、历史价差法、协整模型的计算都比较简单，而 GARCH 模型的建立相对复杂，模型的时效性要差一些。
- ◆ 从模型本身的特点来看，实际价差基本不会回复到公式计算出的理论价差，按持有成本模型进行套利操作不具有可操作性。历史价差法并不能保证样本外数据的均值回复性，同样协整模型也只能保证样本内数据的回复性。而 GARCH 模型较好地刻画了数据的异方差特征，但样本外预测不精确。
- ◆ 从套利次数来看，统计套利的三种方法中，历史价差法套利的次数最少，GARCH 模型法套利的次数最多，由于存在交易成本，反而历史价差法的收益最高，GARCH 模型法的收益最低。
- ◆ 从套利交易的成本来看，当每次套利操作的总成本（包括交易成本、冲击成本）大于 1 个点时，所有的套利模型均亏损。所以，成本高时选择历史价差法更为有利，而成本低时选择 GARCH 模型法更为有利。
- ◆ 由于与其他的期货品种相比，股指期货的套利参与者众多，套利的操作越来越难，套利的机会也越来越少。本文的研究对于其他期货品种的套利机会的挖掘和分析亦有一定的指导作用。

目录

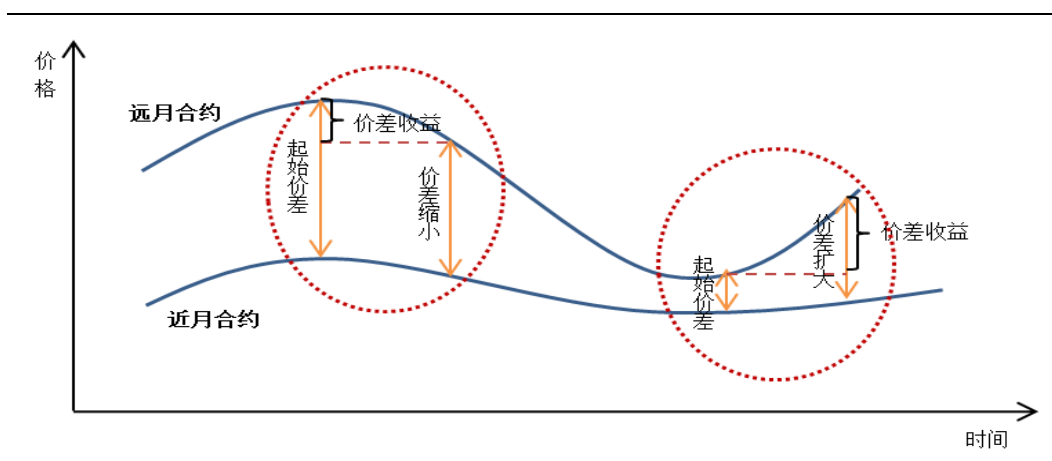
-
- 一．跨期套利的定义
 - 二．基于持有成本模型的跨期套利
 - 三．基于统计套利模型的跨期套利
 - 四．跨期套利模型的比较分析
-

股指期货自2010年4月推出已经三年多，股指期货的套利机会自上市初始就是市场关注的焦点，利用持有成本模型可以计算出两份合约间的理论价值关系，但由于跨期价差不同于期现价差，不一定会随着交割日的临近而收敛，按该理论价值关系进行跨期套利操作的风险很大。而按照预期价差走势的方式来建立套利头寸，在主观性的影响下这种方法局限性也很大。所以我们尝试用统计套利（Statistical Arbitrage）的方法发现价差的稳定性以及变量间的长期均衡关系，用实际的价差与数量模型所预测的价差进行对比，制定统计方法下相对客观的跨期套利策略。本文比较两类不同的跨期套利方法及其收益效果：一种是基于持有成本模型；一种是基于统计套利。

一、跨期套利的定义

跨期套利是指同一投资者以赚取差价为目的，在同一期货品种的不同合约月份建立数量相等、方向相反的交易部位，并以对冲或交割方式结束交易的一种操作方式，因此跨期套利所关注的是合约与合约之间的价差。从原理上来看，当合约间价差出现过高或过低时，相应地卖出或买入价差，直至价差回归合理的均衡价差。

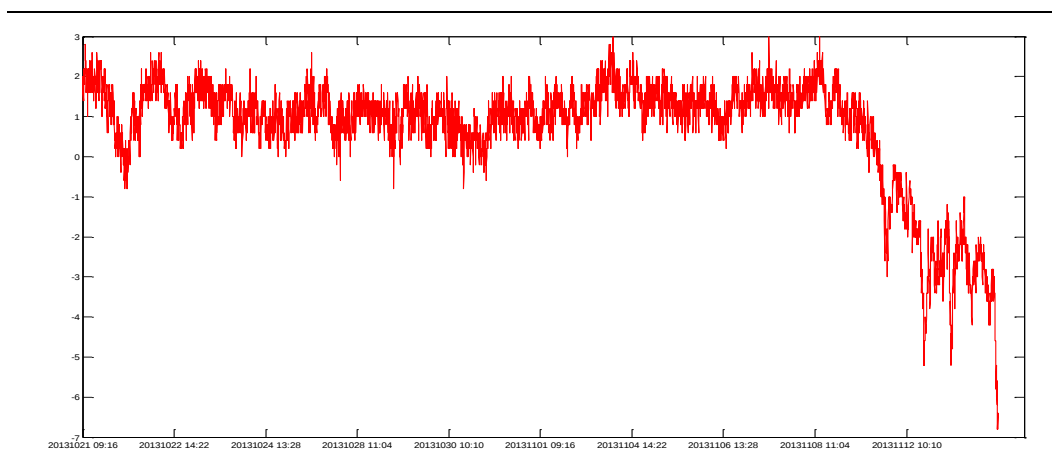
图 1 跨期套利原理图



相比期现套利来说，跨期套利并不是无风险套利，期现套利在合约到期时期现是一定会收敛的，但是跨期价差从理论上来说，可能会随着股指的上升而上升，实际中也可能跟期指投资者对未来中期和短期股指预期的不同而有所变化，因此价差的稳定性需要考察。

本文选取了 IF1311 合约作为主力合约期间的 IF1311 和 IF1312 合约的 1 分钟收盘价数据，即 2013 年 10 月 21 日到 2013 年 11 月 13 日（11 月 15 日为 IF1311 合约的交割日，11 月 14 日主力合约已经转换为 IF1312），图 2 显示的是 IF1312-IF1311 的价差，可以看出价差大多数时间稳定在 0.75 点左右，并不会随着交割日的临近而逐渐收敛。

图 2 IF1312-IF1311 价差图



数据来源：WIND 数据整理：国元期货

二、基于持有成本模型的跨期套利

1. 基于持有成本模型的跨期套利无套利区间

对于同一个标的资产，交易所会设计出不同交割月份的期货合约，对于不同交割月份的股指期货合约理论价格可以表示为：

$$F_t = S e^{(r-d)(T_1-t)} \quad (1)$$

其中： F_t 表示股指期货的时间 t 时的价格； S 表示现货指数在时间 t 时的价格， r 表示以连续复利计算的无风险利率， d 表示股息收益率； T_1 表示期货合约到期时间(年)； t 表示现在的时间(年)。

在假设无风险利率和股息率相同的情况下，任意两张合约之间的理论价格可以表示为：

$$F_i = S e^{(r-d)(T_i-t)} \quad (2)$$

$$F_j = S e^{(r-d)(T_j-t)} \quad (3)$$

由此可以得到两张期货合约之间的价格关系为：

$$F_j = F_i e^{(r-d)(T_j-T_i)} \quad (4)$$

当 $F_j > F_i e^{(r-d)(T_j-T_i)}$ 时，可以执行卖出套利策略；当 $F_j < F_i e^{(r-d)(T_j-T_i)}$ 时，可以执行买入套利策略。

当考虑交易成本时，无论是卖出套利还是买入套利，近远月合约的价差收益必须高于交易成本，才能进行套利操作。因此，在进行卖出套利时，近月远月合约的情况必须满足：

$$\text{Spread} = F_j - F_i > \text{理论价差} + \text{Cost} \quad (5)$$

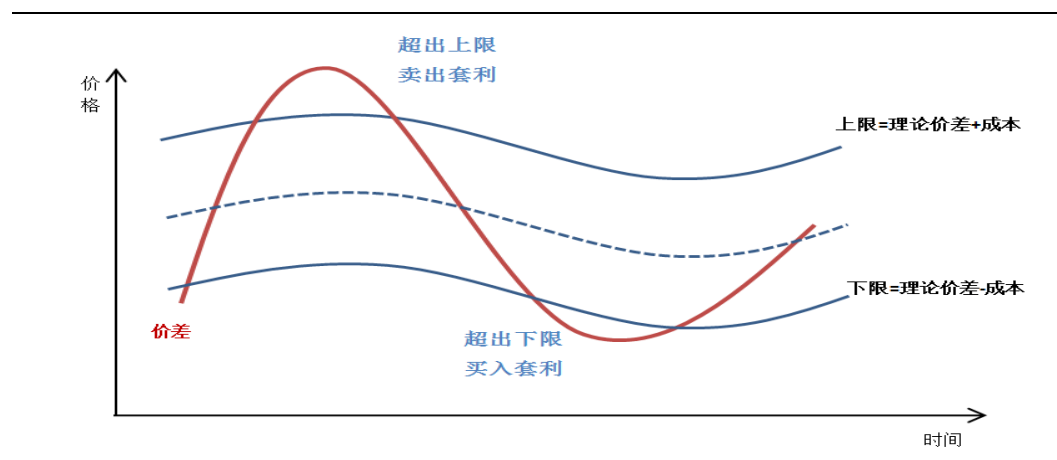
同理，在进行买入套利时应满足：

$$\text{Spread} = F_j - F_i < \text{理论价差} - \text{Cost} \quad (6)$$

其中，理论价差 = $F_i [e^{(r-d)(T_j-T_i)} - 1]$ ，Cost为进行跨期套利时期货交易成本，按照建仓时的期

指价格估算，主要包括期货开平仓的交易成本和冲击成本。由于冲击成本受合约活跃程度、下单手数等多种因素的影响，这里我们暂时先不考虑冲击成本，本文最后就对冲击成本对收益的影响将有另外的说明。

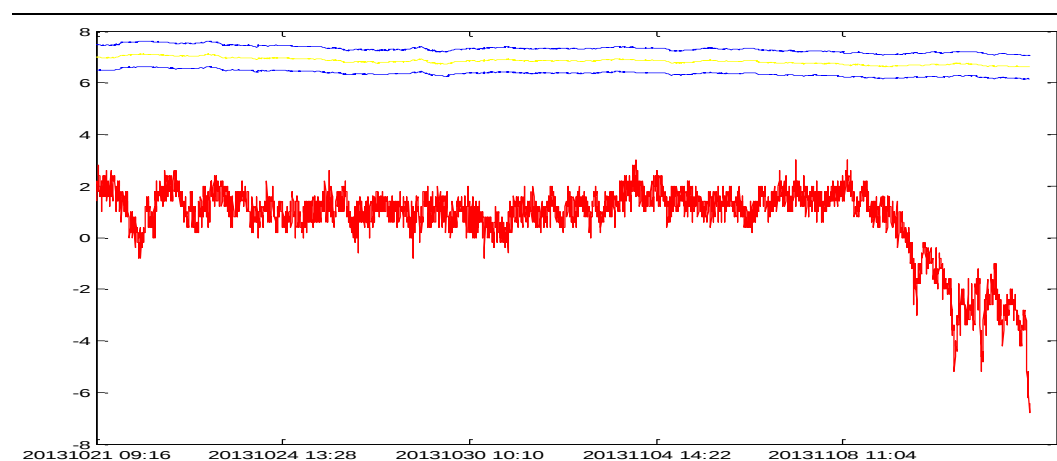
图 3 基于持有成本模型的无套利区间



2. 基于持有成本模型的实证研究

图 4 显示的是2013年10月21日到2013年11月13日的IF1311与IF1312合约的价差及按公式（5）、（6）计算的无套利区间上下限。可以看到由于存在资金成本（设利率 $r=3\%$ ），理论价差整体上移，使得价差一直都处于基于持有成本模型计算的下限之下，无法依据上下限进行套利操作。

图 4 基于持有成本模型的价差图



注：参数 $r=3\%$, $d=0$, 交易成本为成交额的 0.5% , 冲击成本暂不考虑。

数据来源：WIND 数据整理：国元期货

三、基于统计套利模型的跨期套利

第二部分中我们介绍了利用持有成本的分析方法来寻找跨期套利机会。这种方法比较直观，但由于该方法可操作性不强。而且由于资金成本高等原因使得几乎不存在套利。所以，以下我们考虑基于统计

模型的跨期套利。

1. 统计套利的定义及原理

统计套利是基于两个或多个相关性较高的资产组合，如果在未来继续保持这种相关性，那么一旦两者之间出现了背离的走势，则这种背离在未来将会得到纠正，从而产生套利的机会。对统计套利的实践来讲，如果两者间出现背离，那么应该买进表现相对较差的资产而卖出表现相对较好的资产；当未来两者间的背离得到纠正时，反向平仓。其原理背后的基本概念还是均值回复（即资产价格将会回复到它的均值，或者说资产价格序列平稳）。若价差序列是平稳的，那么可以构造统计套利的交易策略，即判断价差是否偏离了长期均值。

2. 统计套利的简单步骤

第一步，按照不同的方法计算出价差（或用残差表示价差），将价差序列 $spread$ 标准化处理：

$$STspread = (spread - \mu) / \sigma \quad (7)$$

其中 μ 为价差的均值， σ 为价差的标准差。取乘数 $k1$ 、 $k2$ 、 $k3$ ，通过历史测试，确定套利区间。其中 $k1$ 为开仓界限， $k2$ 为平仓界限， $k3$ 为止损界限。

第二步，制定套利策略。假设近月合约为 $F1$ ，远月合约为 $F2$ ：

- 当 $STspread > k1$ 时，进行卖出套利，即开仓卖出一张 $F2$ 合约，买入一张 $F1$ 合约；
- 当 $STspread < -k1$ 时，进行买入套利，即开仓买入一张 $F2$ 合约，卖出一张 $F1$ 合约；
- 开仓后，若 $STspread < k2$ ，卖出套利平仓；若 $STspread > -k2$ ，买入套利平仓。
- 开仓后，若 $|STspread|$ 超过标准差过多使得资金紧张，即 $|STspread| > k3$ ，马上止损，由于平稳性理论的保证，在资金足够的条件下可适当扩大止损参数 $k3$ 。

- 当近期合约不再是主力合约，无论价差是否回落，反向对冲了结套利头寸，不进行展期操作。

对于第一步中价差的计算，我们采取三种不同的方法：

- 1) 历史价差法： $spread = F2 - F1$ 。
- 2) 协整模型法：用线性回归方程 $F_{2t} = a + bF_{1t} + \epsilon_t$ 的残差作为 $spread$ 。
- 3) GARCH模型法：将金融时间序列的条件异方差考虑在内，对 $F2-F1$ 建立GARCH模型，以GARCH模型的残差作为 $spread$ 。

3. 统计套利的实证

(1) 历史价差法

历史价差法的核心是选取计算价差均值和标准差的历史滞后时间 n 。一旦 n 值选定，我们可以便可以滚动地计算出 n 分钟的价差均值 μ 和标准差 σ ，进而将 $spread$ 进行标准化处理。假定每次套利的成本固定

为0.5点（每次交易成本为0.125点，开平仓共 $0.125 \times 2 \times 2 = 0.5$ 点，下文中的成本均以此为准），我们按表1中的参数范围进行搭配组合，计算不同参数组合下的套利收益情况。

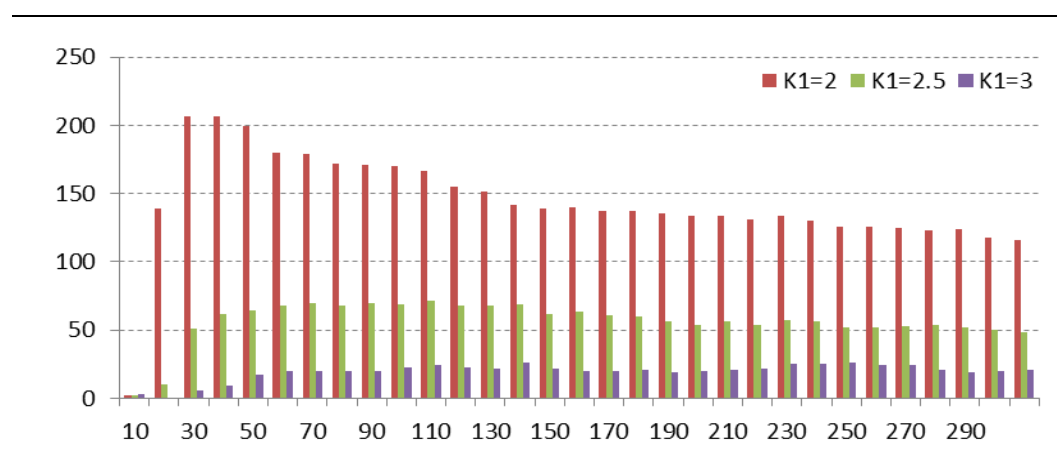
表 1 参数范围

参数	取值范围	参数	取值范围
n	10~300 的整十数	K2	0, 0.5
K1	2, 2.5, 3	K3	3, 3.5, 4, 4.5, 5

数据来源：国元期货

经检验，套利次数会大致随着历史滞后时间n的增加而减少，而K1每增加0.5，套利次数则成倍减少。以n=90为例，K1=2时不同K2、K3取值下的平均套利次数为170次，而K1=2.5、K1=3时的平均套利次数分别为39次、23次。

图 5 参数 n、K1 与套利次数的关系



数据来源：WIND 数据整理：国元期货

我们选择较优的参数组合：n=90，K1=2，K2=0，K3=5。具体的操作为：每次以近90分钟的一分钟收盘价数据计算 μ 和 σ ，每向前一分钟，去掉最早的一个价格，加上最新的一个价格，进而将每次计算的周期固定为90分钟，从而实现滚动向前计算STspread。

具体策略为：

- 当STspread>2时，进行卖出套利；当STspread<-2时，进行买入套利。
- 开仓后，若STspread<0，卖出套利平仓；若STspread>0，买入套利平仓。
- 若STspread>5，卖出套利止损平仓；若STspread<-5，买入套利止损平仓。

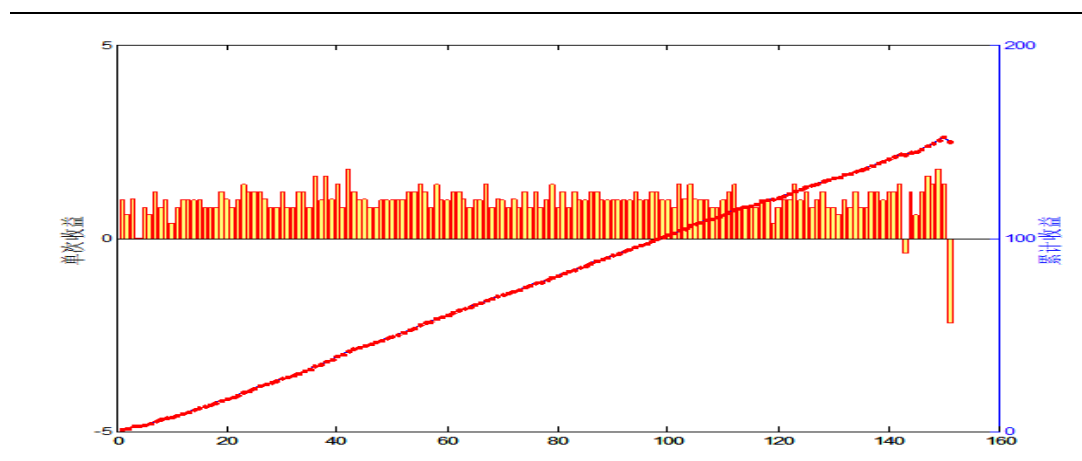
表 2列出了按历史价差法进行套利的统计结果。图 6显示的是套利的单次盈利及累计盈利图。图 7显示的是n=90，K1=2，K2=0，K3=5时的标准化价差及开仓、平仓、止损界限，并标注出了开平仓点。

表 2 历史价差法跨期套利统计结果分析

样本数量	4896	毛盈利（点）	150.2	一次套利所需的最长时间（分钟）	84
卖出套利次数	66	总盈利（扣除交易成本）（点）	74.7	一次套利所需的最短时间（分钟）	1
买入套利次数	85	平均盈利（扣除交易成本）（点）	0.49	一次套利平均需要的时间（分钟）	9.68
跨期套利总次数	151	年化收益率	136.33%	套利持仓时间占总交易时间比例	27.07%

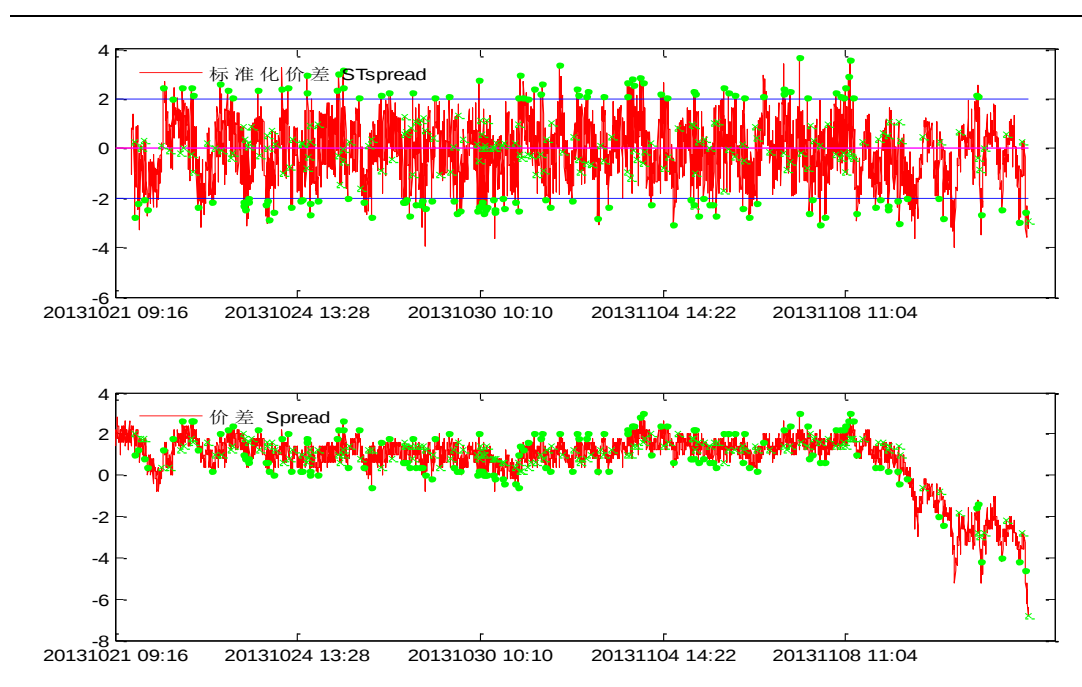
数据来源：WIND 数据整理：国元期货

图 6 历史价差法跨期套利单次盈利及累计盈利图



数据来源：WIND 数据整理：国元期货

图 7 历史价差法跨期套利机会图



注：图中绿点表示开仓点，绿叉表示平仓点 数据来源：WIND 数据整理：国元期货

(2) 协整模型法

由于不同合约的走势都是基于对未来标的指数的预期产生的，除了持有成本带来的合约价差外还有一些非合理的因素，从长期来看同一标的的各合约价格之间存在着这样一种平稳关系。协整概念便是处理非平稳时间序列的较好统计方法。

基于协整模型的统计套利具体步骤如下：

第一步，首先考虑近月合约F1与远月合约F2两个总体的线性相关性。如果总体之间没有显著的相关性，很难以统计套利的方法制定策略；反过来说，虽然较高的相关性并不代表稳定的协整关系，但这时序列间协整的可能性会更大，统计套利的可行性也就更大。若达到我们规定的相关性，则继续下面的步骤。

第二步，金融时间序列大多是不平稳的，即价格序列本身并不是均值回复的，它们往往是一阶或二阶单整序列。我们要验证的就是两个时间序列F1、F2是否同阶单整，即对序列作单位根检验，若经过相同的差分次数后（也可以是0次，即两序列都平稳），两序列同时通过了单位根检验，则它们是同阶单整的。

第三步，对同阶单整序列F1、F2作线性回归：

$$F_{2t} = a + bF_{1t} + \varepsilon_t \quad (8)$$

由于有了第一步的铺垫，它们可以达到一定的相关性。取出线性回归的残差，进行单位根检验，若通过，说明残差是均值回复的。

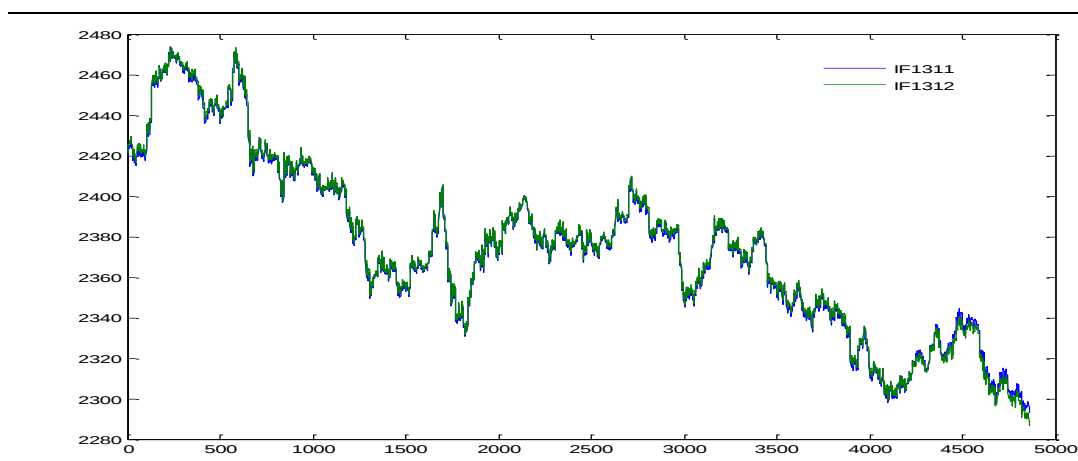
第四步，将平稳的残差序列 $\{\varepsilon_t\}$ 作为价差序列spread，按照公式（7）标准化为STspread，设定乘数k1、k2、k3，然后按照上一部分中的统计套利步骤进行开平仓。

我们代入2013年10月21日到2013年11月14日的IF1311与IF1312合约的1分钟收盘价数据，按照以上的步骤进行协整检验。

a) F1、F2的线性相关性

IF1311合约与IF1312合约的1分钟收盘价的相关系数高达0.9996，从图 8来看，两条线几乎重合，也印证了两个合约的价格序列具有高度相关性。

图 8 IF1311 合约与 IF1312 合约的 1 分钟收盘价



数据来源：WIND 数据整理：国元期货

b) 单位根检验

对IF1311、IF1312序列进行单位根检验，从检验结果（如表 3所示）来看，IF1311和IF1312序列都含有单位根，但两个序列的一阶差分是平稳的，即都是一阶单整，可以进行协整分析。

表 3 IF1311 和 IF1312 的单位根检验结果

变量	H	P	ADF	临界值（5%）	是否平稳
IF1311	0	0.232	-1.1475	-1.9416	否
IF1312	0	0.198	-1.2402	-1.9416	否
$\Delta IF1311$	1	0.001	-67.997	-1.9416	是
$\Delta IF1312$	1	0.001	-68.843	-1.9416	是

数据来源：WIND 数据整理：国元期货

c) 对同阶单整序列F1、F2作线性回归

利用一分钟高频价格数据对IF1311、IF1312进行回归可得：

$$IF1312 = -32.5612 + 1.0140 * IF1311 + resid \quad (9)$$

$$(R^2=0.9992, F=5870322.27, P=0)$$

模型的残差单位根检验结果如下：

表 4 残差的单位根检验结果

变量	H	P	ADF	临界值（5%）	是否平稳
resid	1	0.001	-11.35	-1.9416	是

数据来源：WIND 数据整理：国元期货

由表 4可知，残差的单位根检验表明其基本上是平稳的，所以IF1312和IF1311存在协整关系。

以上是对两个合约序列总体的协整关系的检验。在实际操作过程中，样本只是一段时间。与历史价差法类似，我们也要设定计算价差均值和标准差的周期n，同时n也是进行回归分析时的样本数量。我们同样采用历史价差法中的参数组合：n=90，K1=2，K2=0，K3=5。

具体的操作为：每次以近90分钟的一分钟收盘价数据为样本进行回归，得到回归方程 $IF1312 = a + b * IF1311 + resid$ ，代入IF1312、IF1311的1分钟收盘价得到残差序列 $resid = IF1312 - a - b * IF1311$ ，用残差resid代表价差spread，按照公式（7）进行标准化处理。具体策略同上。

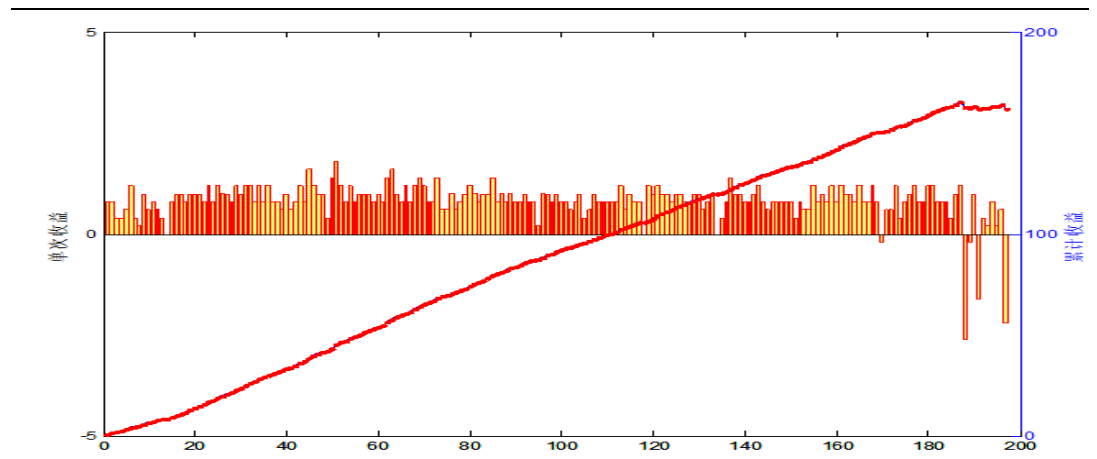
根据统计计算，我们得到协整模型法跨期套利的统计结果如表 5所示。图 9显示的是协整模型法跨期套利的单次盈利及累计盈利图。图 10显示的是n=90，K1=2，K2=0，K3=5时的标准化价差及开仓、平仓、止损界限，并标注出了开平仓点。

表 5 协整模型法跨期套利统计结果分析

样本数量	4896	毛盈利（点）	161.8	一次套利所需的最长时间（分钟）	68
卖出套利次数	90	总盈利（扣除交易成本）（点）	63.3	一次套利所需的最短时间（分钟）	1
买入套利次数	107	平均盈利（扣除交易成本）（点）	0.32	一次套利平均需要的时间（分钟）	7.56
跨期套利总次数	197	年化收益率	115.52%	套利持仓时间占总交易时间比例	27.57%

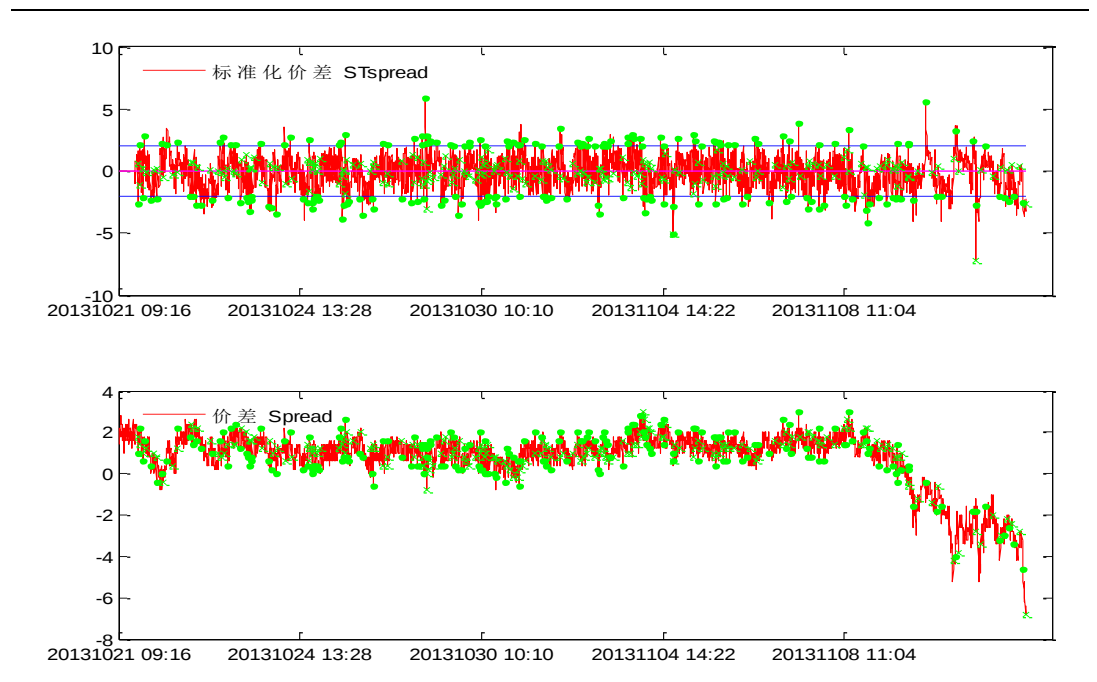
数据来源：WIND 数据整理：国元期货

图 9 协整模型法跨期套利单次盈利及累计盈利图



数据来源：WIND 数据整理：国元期货

图 10 协整模型法跨期套利机会图



数据来源：WIND 数据整理：国元期货

由于公式（8）中的b在大多数情况下不是整数，并且在实时交易中不能指定一个确定的比例进行交易（确定的比例在大多数情况下无法达到统计残差的均方误差最小化），我们在编写程序时采用将1: b约简成两个整数相比的方法后，再进行期货交易，股指期货的跨期套利中是1: 1。这种做法保证了统计套利策略的完整性，缺点是影响了整体策略的收益，这也是采用协整模型法所要面对的一大问题。

（3） GARCH 模型法

以上是利用协整模型的方法，采取了较简单的线性回归方程建立模型，并以残差resid自身的标准差对残差序列进行标准化处理。但实际上，resid序列自身的方差并非恒定不变，和其他大多数金融时间序列一样，resid也具有时变方差特性，需要对价差的方差进行预测，下面建立GARCH(1,1)模型对样本外价差的时变方差进行估计。

在GARCH模型中，要考虑两个不同的设定：一个是条件均值，另一个是条件方差。标准的GARCH(1,1)模型为：

$$Y_t = X_t \gamma + u_t, t = 1, 2, \dots, T$$

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (10)$$

其中：X_t是1×(k+1)维外生变量向量，γ是(k+1) ×1维系数向量。

我们以2013年10月21日至建模时间点的IF1312-IF1311价差的1分钟数据作为样本内数据建立AR(4)-GARCH模型：

$$\text{Spread}_t = a + b_1 \text{Spread}_{t-1} + b_2 \text{Spread}_{t-2} + b_3 \text{Spread}_{t-3} + b_4 \text{Spread}_{t-4} + \varepsilon_t$$

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (11)$$

随着时间的推移，样本数据不断的增加，由于盘中无法及时建立GARCH模型，设定为每半个交易日对模型进行更新。

将样本外数据代入GARCH(1,1)模型，计算残差及标准化残差：

$$\text{resid} = \text{Spread}_t - (a + b_1 \text{Spread}_{t-1} + b_2 \text{Spread}_{t-2} + b_3 \text{Spread}_{t-3} + b_4 \text{Spread}_{t-4}) \quad (12)$$

$$\text{STspread} = (\text{resid} - \mu) / \sigma \quad (13)$$

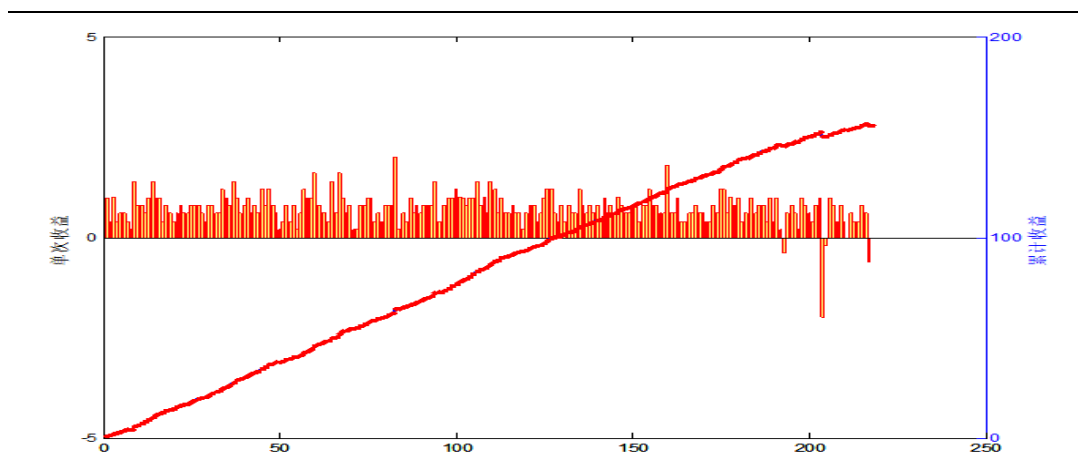
交易策略的制定我们仍然采用参数组合：n=90，K1=2，K2=0，K3=5。按GARCH模型法进行套利的统计结果如表 6所示，套利的单次盈利及累计盈利如图 11所示，具体的开平仓点如图 12所示。

表 6 GARCH 模型法跨期套利统计结果分析

样本数量	4896	毛盈利（点）	156	一次套利所需的最长时间（分钟）	28
卖出套利次数	82	总盈利（扣除交易成本）（点）	47	一次套利所需的最短时间（分钟）	1
买入套利次数	136	平均盈利（扣除交易成本）（点）	0.22	一次套利平均需要的时间（分钟）	2.62
跨期套利总次数	218	年化收益率	85.77%	套利持仓时间占总交易时间比例	10.59%

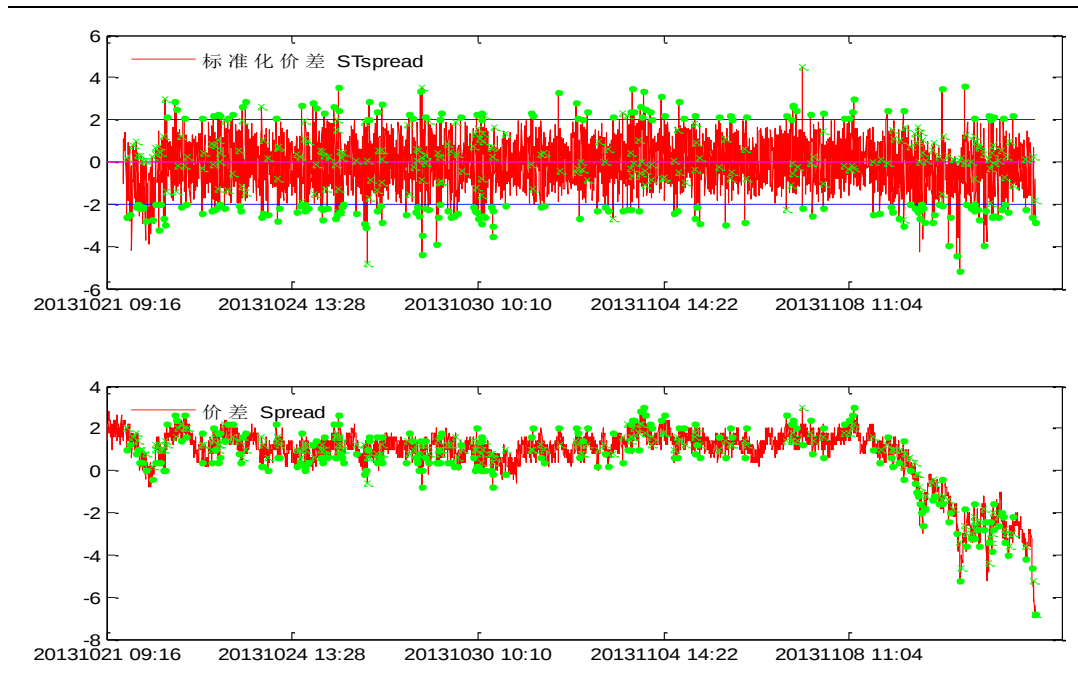
数据来源：WIND 数据整理：国元期货

图 11 GARCH 模型法跨期套利单次盈利及累计盈利图



数据来源: WIND 数据整理: 国元期货

图 12 GARCH 模型法跨期套利机会图



数据来源: WIND 数据整理: 国元期货

四、跨期套利模型的比较分析

以上我们分别就持有成本模型、历史价差法、协整模型和GARCH模型的跨期套利机会进行了计算，各种计算方法的特点不同，各有优劣：

从价差的计算方法来看，持有成本模型、历史价差法、协整模型的计算都比较简单，利用程序化交易软件可以实现在盘中实时计算，随着行情不断地更新，而GARCH模型的建立相对复杂，只能在休市时更新模型，模型的时效性要差一些。

从模型本身的特点来看，实际价差基本不会回复到公式计算出的理论价差，按持有成本模型进行套利操作不具有可操作性。历史价差法并不能保证样本外数据的均值回复性，同样协整模型也只能保证样本内数据的回复性。而GARCH模型较好地刻画了数据的异方差特征，但样本外预测不精确。

从套利次数来看，统计套利的三种方法中，历史价差法套利的次数最少（151次），GARCH模型法套利的次数最多（218次），由于存在交易成本，反而历史价差法的收益最高（74.7点），GARCH模型法的收益最低（47点）。

从套利交易的成本来看，上文的分析中只考虑了交易成本（0.5点），计算结果显示当冲击成本大于0.5点，即总成本大于1个点时，所有的套利模型均亏损。所以，根据成本的不同可以选择不同的模型，成本高时选择历史价差法更为有利，而成本低时选择GARCH模型法更为有利。

此外仍有一些值得思考的地方。本文都是以建模时点之前的收盘价为样本数据，再用样本外数据代入求得标准化残差，虽然收益性比样本内建模要差，但更贴近实际情形。由于与其他的期货品种相比，股指期货的套利参与者众多，套利的操作越来越难，套利的机会也越来越少。本文的研究对于其他期货品种的套利机会的挖掘和分析亦有一定的指导作用。

风险提示：期市有风险，入市须谨慎！

免责声明：本报告中的信息均来源于公开可获取性资料，国元期货力求客观公正，但不保证这些信息的准确及完整，也不保证这些信息未经任何更新，更不保证我们的建议或意见不会发生任何变更。报告中的信息和我们的意见乃至结论并不构成所述品种交易的出价或征价，投资者据此进行的投资行为与我公司和作者无关。

分析师简介：

戴妍，国元期货金融工程研究员，2010 年起开始从事金融工程量化研究工作，自 2012 年起进入期货行业，主要致力于趋势型、套利型程序化交易策略的构建。曾在《期货日报》发表策略文章。2012 年参与十大研发团队比赛，所在农产品团队获得总成绩第十三名的荣誉。从业资格证号 F0282161。

电话：021-68880376

邮箱：daiyan@guoyuanqh.com