基于GARCH模型的股指期货跨期套利研究

Study on Inter-temporal Arbitrage from Stock Index Futures Based on GARCH Model

莫海军
（中信建投期货有限公司　重庆　400014）

摘要：针对股指期货当月合约与下月合约间的价差存在波动聚集的现象，为了改善跨期套利效果，基于股指期货合约间的比值关系，本文利用GARCH（1,1）模型建立了股指期货跨期套利模型。实证分析表明，该模型在跨期套利中运用较为有效。若价差波动较小，该模型能够有效捕捉到价差异动点；若价差波动较为剧烈，套利条件会迅速得到放松，降低风险。

关键词：股指期货　GARCH模型　跨期套利

一、引言

沪深300股指期货上市以来，基于股指期货合约的套利交易以其风险小、收益稳定等特点迅速被广大投资者接受，交易方式主要分为期现套利和跨期套利。但是，由于股指期货市场发展迅速，对于期现套利而言，无论是套利机会还是套利收益均比股指期货上市初期大幅减少。而跨期套利具有操作简单、风险小、机会较多等优点，只要流动性能得到保证，跨期套利必将得到快速发展。股指期货的四个合约中，只有当月合约和下月合约的流动性较好，本文的研究主要以这两个合约为基础展开。

常见的跨期套利策略是基于一定的时间频率（1分钟、5分钟或10分钟），取一定长度的历史数据，统计得出价差的概率分布，通过该价差分布来判断目前价差是处于过大、过小还是合理水平。若认为目前价差过大或者过小，投资者可在期货市场建立相应头寸，在价差回归合理的过程中获利。对于这种操作策略，最致命的问题是价差分布不稳定，统计得到的分布情况依赖于样本统计区间。因为分布不稳定，一旦市场出现异常波动，这种策略无法得到及时调整，必将导致跨期套利出现亏损。以2010年为例，统计4月至9月的价差，我们会认为价差超过50点已经非常高，仍会维持原来的开平仓策略。事实上，后来价差一度超过100点。日内数据也是一样，一旦日内出现大涨或大跌，价差很可能出现急剧变化，这就要求我们的模型能及时适应市场变化。

事实上，与大多数金融数据一样，股指期货合约间价差也呈现出波动性聚集（volatility clustering）的特征，即在相当长一段时间内大幅波动，然后又会在下一段时间内保持相对稳定。这就要求我们的模型能随着市场的变化得到及时调整。

二、ARCH模型与GARCH模型简介

为了刻画变量的波动聚集性，Engle（1982）提出建立条件异方差模型即ARCH（Autoregressive Conditional Heteroskedasticity）模型，并用来对条件异方差进行预测。P阶ARCH模型包括均值方程和条件方差方程：

$$\begin{cases} y_t = \theta x_t + \varepsilon_t \\ \delta_t^2 = \mathrm{var}(\varepsilon_t | \Omega_{t-1}) = c + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots\cdots + \alpha_p \varepsilon_{t-p}^2 \end{cases}$$

其中，Ω_{t-1}表示t–1时刻所有可得信息的集合。条件方差方程的含义即误差项ε_t的方差δ_t^2依赖于前p个时期的误差平方。即若前期波动较大使得残差较大，那么当期的波动性也会增加。

Bollerslev（1986）在ARCH模型的基础上提出了GARCH（Generalized ARCH）模型，其形式为：

$$\begin{cases} y_t = \theta x_t + \varepsilon_t \\ \varepsilon_t \sim N(0, \delta_t^2) \\ \delta_t^2 = \mathrm{var}(\varepsilon_t | \Omega_{t-1}) = c + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{t=j}^{q} \beta_j \delta_{t-j}^2 \end{cases}$$

即条件方差由常数项、前期残差平方以及前期条件方差构成。其中：$c \geq 0$，$\alpha_i \geq 0, \beta_j \geq 0, \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$。这样，模型能很好地解释波动聚集现象。

三、股指期货合约间价格关系分析

股指期货跨期套利的逻辑基础是股指期货合约价格之间应保持相对稳定的关系。这种稳定的关系可以是价差关系或者比值关系。若不同合约之间存在稳定的价差关系，在进行跨期套利时，我们可对其价差进行分析。若合约之间存在稳定的比值关系，那么就应该对其回归后的残差进行分析。

股指期货每个合约均有一定的生命周期，当合约到期后，该合约就会退出市场。为了系统分析股指期货当月合约和下月合约之间的价格关系，本文使用了当月连续合约和下月连续合约进行计算，时间区间为2010年4月16日至2012年5月16日，当月连续合约由每个交易日在交易的合约中离交割日最近的合约连接构成，下月连续合约由每个交易日在交易的合约中次月交割的合约连接构成。

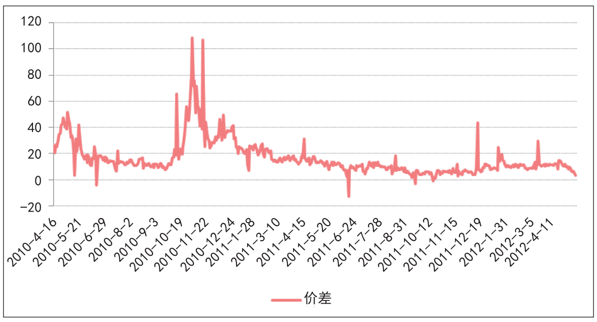
（一）合约间价差关系分析

对合约间历史价差进行单位根检验，结果见表1，显示在5%的显著性水平下，当月合约与下月合约之间的价差不平稳。

表1 价差单位根检验

	ADF值	检验形式	P值	结论
回归残差	–3.40	（C,T,17）	0.052	不平稳

从图1可以看出，价差呈现出明显的波动聚集性，在2010年4月至5月以及2010年10月至11月这两段时间内表现最明显。



数据来源：WIND 资讯

图1 下月连续合约与当月连续合约价差走势

（二）合约间比值关系分析

和其他金融时间序列一样，当月合约和下月合约价格本身并不平稳（表2），但它们的一阶差分序列是平稳的（表3），也就是说它们是一阶单整序列。

表2 当月、下月连续合约价格单位根检验

	ADF值	检验形式	P值	结论
当月连续	−2.13	（C,T,0）	0.52	不平稳
下月连续	−1.85	（C,T,1）	0.68	不平稳

表3 当月、下月连续合约价格一阶差分单位根检验

	ADF值	检验形式	P值	结论
当月连续一阶差分	−24.19	(C,T,0)	0	平稳
下月连续一阶差分	−24.58	（C,T,0）	0	平稳

由表3可见，在进行一阶差分处理后，当月和下月连续合约价格均在很小的显著性水平下（1%）拒绝含有单位根的原假设，表明两个价格数据均为一阶单整序列。对其进行协整分析发现（如公式1和表4所示）他们之间保持长期平稳的关系。

$$\hat{y} = 1.0059x \ (adjustedR^2 = 0.9981) \qquad (1)$$

表4 协整检验

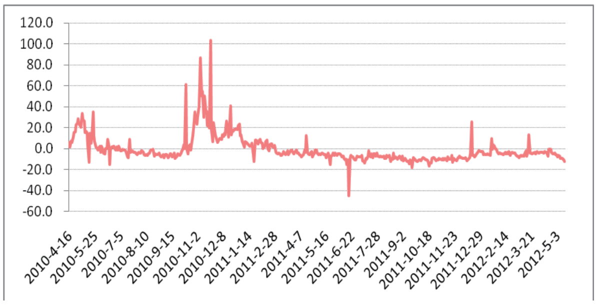
	ADF值	检验形式	P值	结论
回归残差	−3.47	（C,T,17）	0.044	平稳

利用ARCH LM方法对残差进行ARCH效应检验，滞后阶数为3的检验结果见表5。从表5的结果可知，以P值为0拒绝原假设，这说明公式（1）的残差序列存在ARCH效应。

表5 ARCH LM检验结果（日数据）

F统计量	32.1668	P值	0.000
T*R统计量	81.4095	P值	0.000

从图2也可以看出，在比值关系上，残差仍表现出波动聚集性。我们利用股指期货上市以来的数据，对1分钟、5分钟和15分钟数据进行类似的回归分析（表6），残差均存在ARCH效应。



数据来源：中信建投期货

图2 下月连续合约与当月连续合约回归残差

表6 ARCH LM检验结果（高频数据）

1min	F统计量	4017806	P值	0.0000
	T*R统计量	134934	P值	0.0000
5min	F统计量	255034	P值	0.0000
	T*R统计量	27079	P值	0.0000
15min	F统计量	37431	P值	0.0000
	T*R统计量	8351	P值	0.0000

（三）ARCH效应在跨期套利中的危害

跨期套利的前提是价差或残差服从稳定的分布。若价差或残差表现出波动聚集性，这意味着高价差或低价差会连续出现。所以当价差持续维持高位或者持续维持低位时我们需对套利策略进行调整。当价差持续保持在较低水平时，若不对策略进行调整，很可能长期捕捉不到套利机会。当价差持续保持在较高水平时，需对套利临界值进行调整，否则模型将持续报出无操作意义的开仓信号。

四、GARCH模型在跨期套利中的应用

（一）GARCH模型的建立

利用GARCH模型可消除残差的群集波动性。在GARCH模型中，残差服从正态分布，只是正态分布的方差是时变的。可以利

用以前的残差平方和条件方差来预测下一时刻残差的条件方差。在得到残差条件方差的预测值后，我们就可衡量该残差是否异常。

分析发现，GARCH（1,1）模型在大多数时候可消除下月合约与当月合约回归方程残差的波动聚集性。虽然我们无法用单个方程对所有股指期货数据进行很好的拟合，但在单个交易日，GARCH（1,1）模型能够做到这一点。我们主要用1分钟数据对其进行分析。当月合约记为if_{1t}，下月合约记为if_{2t}。以2012年5月2日为例，利用OLS对IF1206合约1分钟数据和IF1205合约1分钟数据回归，得到的结果如下：

$$if_{2t}=1.0039\,if_{1t}+\hat{\varepsilon}_t \qquad (2)$$

对（2）式进行ARCH LM检验，得到滞后阶数为3时的ARCH LM检验结果，见表7。

表7 ARCH LM检验结果

F统计量	26.275	P值	0.000
$T\times R^2$ 统计量	61.670	P值	0.000

表7表明公式（2）残差存在ARCH效应。利用GARCH（1,1）模型重新估计，结果如下：

均值方程：

$$if_{2t}=1.0039\,if_{1t}+\hat{\varepsilon}_t \qquad (3)$$

z =64409

方差方程：

$$\hat{\delta}_t^2=0.0326+0.1593\,\hat{\varepsilon}_{t-1}^2+0.7554\,\hat{\delta}_{t-1}^2 \qquad (4)$$

z =（1.07）　（2.22）　（6.01）

方差方程中的ARCH项和GARCH项的系数都是统计显著的，并且系数之和小于1，满足GARCH模型的参数约束条件。由于系数之和非常接近于1，表明条件方差所受冲击是持久的。对（3）式残差序列进行ARCH LM检验，滞后阶数为3的统计结果见表8。

表8 ARCH LM检验结果

F统计量	0.596	P值	0.618
$T\times R^2$ 统计量	1.802	P值	0.615

统计结果接受原假设，认为该残差序列已不存在ARCH效应，这说明GARCH（1,1）模型消除了（2）式残差的ARCH效应。

对2012年5月2日至2012年5月16日的1分钟数据进行类似分析，可以发现，即使市场并未出现股指期货上市初期或者2010年10月那样市场剧烈波动的情况，但ARCH效应仍然非常显著，而GARCH（1,1）模型能对其进行消除。对OLS模型的回归残差和GARCH（1,1）模型回归残差进行ARCH LM检验，结果分布见表9和表10。

表9 OLS模型回归残差的ARCH LM检验结果

5月3日	F统计量	12.97	P值	0.000
	T*R2统计量	12.47	P值	0.000
5月4日	F统计量	19.860	P值	0.000
	T*R2统计量	49.369	P值	0.000
5月7日	F统计量	14.326	P值	0.000
	T*R2统计量	37.520	P值	0.000
5月8日	F统计量	51.101	P值	0.000
	T*R2统计量	98.454	P值	0.000
5月9日	F统计量	51.101	P值	0.000
	T*R2统计量	98.454	P值	0.000
5月10日	F统计量	20.927	P值	0.000
	T*R2统计量	51.487	P值	0.000
5月11日	F统计量	35.101	P值	0.000
	T*R2统计量	76.417	P值	0.000
5月14日	F统计量	42.957	P值	0.000
	T*R2统计量	87.910	P值	0.000
5月15日	F统计量	20.536	P值	0.000
	T*R2统计量	50.708	P值	0.000
5月16日	F统计量	140.848	P值	0.000
	T*R2统计量	164.945	P值	0.000

表10 GARCH（1,1）模型回归残差的ARCH LM检验结果

5月3日	F统计量	1.042	P值	0.374
	T*R2统计量	3.137	P值	0.371
5月4日	F统计量	0.685	P值	0.562
	T*R2统计量	2.070	P值	0.558
5月7日	F统计量	1.832	P值	0.142
	T*R2统计量	5.466	P值	0.141
5月8日	F统计量	0.336	P值	0.800
	T*R2统计量	1.018	P值	0.797
5月9日	F统计量	0.336	P值	0.800
	T*R2统计量	1.018	P值	0.797
5月10日	F统计量	0.199	P值	0.897
	T*R2统计量	0.604	P值	0.896
5月11日	F统计量	2.187	P值	0.090
	T*R2统计量	6.499	P值	0.090
5月14日	F统计量	0.590	P值	0.622
	T*R2统计量	1.785	P值	0.618
5月15日	F统计量	0.988	P值	0.399
	T*R2统计量	2.976	P值	0.395
5月16日	F统计量	0.259	P值	0.855
	T*R2统计量	0.786	P值	0.853

除了5月7日当天数据出现了GARCH项小于0的情况外，其余系数均满足模型条件。可见GARCH（1,1）模型可消除大多数交易日下月合约与当月合约价格回归后残差的ARCH效应。

（二）基于GARCH模型的跨期套利策略制定

在t时刻，我们可以利用标准差δ_t来判断ε_t是否异常。将其标准化，$\mu_t=\dfrac{\varepsilon_t-0}{\delta_t}$，我们可用$\mu_t$来判定$\varepsilon_t$是否极端。若$\mu_t$>2或$\mu_t$<2，我们认为出现这种情况的概率不足2.5%，残差的异常意味着价差出现异常，这是跨期套利建仓的时机。当μ_t的绝对值较小时，比如$|\mu_t|$<0.5，可以认为残差已经正常，也意味着价差回归正常，这是平仓的时机。

与期现套利不同，跨期套利没有价差必然回归正常的制度保证。因此，任何跨期套利策略必须要有止损措施。比如，当$|\mu_t|$大于一定数值时（如3），投资者可考虑离场。从我们跟踪的情况看，这种需要止损的情况较为少见。

（三）模型修正

在应用GARCH（1,1）模型时，我们利用样本内数据来预测样本外的条件方差。为了使模型具有更好的预测能力，理想的情况是利用实时成交数据对方程参数进行滚动估计，但这对系统要求较高。对于一般分析而言，可每隔一定时间长度进行一次估计。有时会出现模型估计参数不符合模型要求的情况，比如ARCH项或者GARCH项系数小于0或者两者之和大于1。在这种情况下，需舍弃本次估计结果，继续使用上一周期所估计的数据。

（四）实证分析

在跟踪分析过程中，每个交易日对模型重估一次，用上一交易日成交数据估算的参数来指导当日交易。在2012年5月2日数据估计的模型参数的基础上，我们对2012年5月3日的跨期套利机会进行分析，结果如图3、图4以及表11所示。

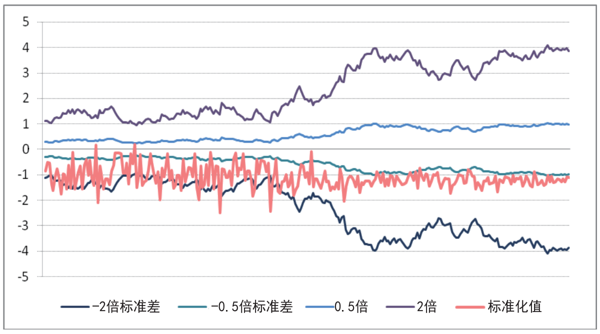


图3 2012年5月3日跨期套利机会分析

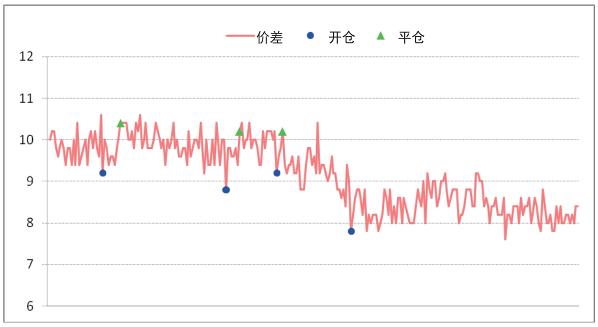


图4 2012年5月3日开平仓时机

表11 高频跨期套利分析结果（2012年5月3日）

	IF1206	IF1205	价差	标准化	操作方式	利润分析
9:43:00	2690.8	2700	9.2	−2.11	开仓	–
9:52:00	2690.2	2700.6	10.4	−0.12	平仓	360
10:46:00	2690.2	2699	8.8	−2.50	开仓	–
10:53:00	2692.6	2702.8	10.2	−0.36	平仓	420
11:12:00	2690	2699.2	9.2	−2.44	开仓	–
11:15:00	2686	2696.2	10.2	−0.38	平仓	300
13:19:00	2690.6	2698.4	7.8	−2.04	开仓	–

从上面的检验结果可知，该模型能有效指导股指期货高频跨期套利。在价差平稳波动的时候，该模型能有效捕捉到价差的异常波动，当价差出现结构性变化时（比如2012年5月3日下午交易时段），也不会频繁报出无操作意义的开平仓信号。

（五）GARCH（1,1）模型与Boll通道

Boll通道是根据标准差原理设计的，因此Boll通道也会随着价格的变化而自动调整。一种思路是，对价差用Boll通道进行分析，一旦价差变大，标准差也会变大，Boll通道变宽，这也能达到GARCH（1,1）模型的效果。

但跟踪研究后发现，由于Boll通道的宽度取决于前面一段历史数据的标准差，这会降低模型反应灵敏度。当价差瞬间放大后，根据前面一段时间计算得到的标准差所受影响很小，即Boll线反应不灵敏。从这点讲，GARCH（1,1）模型无法用Boll通道替代。

五、结论与展望

股指期货当月合约和下月合约价差存在波动聚集性。一般跨期套利模型并未考虑价差的这种特性，这会使得模型无法有效捕捉跨期套利机会。GARCH（1,1）模型能够有效消除股指期货价格之间回归残差的波动聚集性。文中利用GARCH（1,1）模型建立了跨期套利模型，从实证研究结果看，该模型能够很好地捕捉跨期套利机会，并提示合适的平仓机会。

本文仅仅验证了这种操作思路的可行性，对具体模型设计过程中的参数优化并未作深入分析，比如数据频率、开平仓条件等。在跟踪分析中我们也发现GARCH（1,1）模型无法对所有数据进行很好的拟合，未来研究重点将是对现有模型的进一步优化。

参考文献

1、ENGLE, Robert F., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation, Econometrica 50:987−1008.

2、BOLLERSLEV, Tim, 1986. Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31:307−327.

（责任编辑　李泽海）