

文章编号: 1001 - 9081(2013) S1 - 0326 - 03

基于 Black-Scholes 方程的股指期货期现套利模型及交易算法

王 飞, 孙维尧*

(河海大学 商学院, 南京 211100)

(* 通信作者电子邮箱 wy_sun1988@gmail.com)

摘 要: 为了适应股指期货期现套利高频交易需要, 解决传统协整理论不能很好捕捉交易信号的问题, 构建了基于 Black-Scholes 方程的期现套利模型及算法。首先使用修正跟踪误差指标确定交易所交易基金投资组合最优化权重并用广义既约梯度算法求解; 然后构建了基于 Black-Scholes 方程的套利模型和算法交易系统。实证检验显示, 该算法实现了 1.52% 的单日净收益率和 5% 的止损发生率, 且止损发生率与波动性正相关。证明该算法在市场条件稳定的条件下能够很好地捕捉交易信号和控制交易行为, 如果市场发生较大异常波动, 算法的适用性将受到影响。

关键词: 期现套利; 高频交易; 交易算法; 股指期货

中图分类号: TP301.6; F830.91 **文献标志码:** A

Stock index futures arbitrage model based on Black-Scholes equation and corresponding trading algorithm

WANG Fei, SUN Weiyao*

(Business School, Hohai University, Nanjing Jiangsu 211100, China)

Abstract: To meet the practical need of high frequency index future-spot arbitrage and overcome the disadvantage of the arbitrage algorithms based on traditional cointegration theory, which is failing to recognize many trading signals, a new index future-spot arbitrage model and trading algorithm were established. Based on modified tracking error model and Generalized Reduced Gradient (GRG) algorithm, an optimized weighting ETF (Exchange Trade Fund) portfolio was built and then the arbitrage model and the algorithm trading system based on Black-Scholes equation were formulized. The empirical results show that, the algorithm realized a daily net yield of 1.52% and a stop-loss rate of 5%, and the stop-loss rate is positively correlated with the market volatility. It is proved that this algorithm has a good control on recognizing the trading signal and executing trading orders under stable market condition while its practicability will be discounted under the market turmoil.

Key words: futures-spot arbitrage; high frequency trading; trading algorithm; stock index futures

0 引言

高频交易(也称算法交易)是随着高速计算机发展和相关经济理论进步而发展起来的新兴领域。股票指数期货与交易型开放式指数基金(Exchange Traded Fund, ETF)之间的套利算法开发是高频交易研究的重点方向。Cornell 等^[1] 构建的期货持有成本模型为两者之间的套利奠定了理论基础。国外已经有若干学者利用协整理论构建股指期货与 ETF 的套利模型^[2-5], 针对美国标准普尔指数期货和以标准普尔 500 指数存托凭证(Standard & Poor's Depository Receipts, SPDR)的协整关系和套利策略进行了研究; 近年来国内也有学者^[6-8] 将同样的方法应用到了国内市场。然而由于当时国内的股指期货还未上市, 所以使用的是模拟交易数据。总体来看, 前人研究的基本思路都是利用对股指期货与 ETF 的价格时间序列数据进行协整检验, 若存在协整关系则有套利的可能。这类套利方法的问题在于其所发现的是两者价格背离长期关系时的套利机会, 而在高频交易中, 由于所分析的价格数据集中在日内, 数据的随机性较高, 传统的协整检验对这种微观的市场结构进行分析会造成普通最小二乘法(Ordinary Least Squares, OLS)估计有偏、协整检验结果可能有误的问题^[9]。

为了解决上述问题, 本文从 Black-Scholes (B-S) 方程出

发, 构建了基于股指期货和 ETF 的无套利投资组合, 由此推导出极小时间和极小标的指数变动下的价差序列, 最后利用 IF300 和 ETF 组合进行了实证研究并计算了套利收益。

1 基于 B-S 方程的股指期货期现套利模型

1.1 ETF 组合的构建

股指期货期现套利中现货工具的选择应考虑三个方面: 跟踪误差小、交易成本低、流动性好。通常有以下几种选择: 成份股全部复制^[10]、成份股部分复制^[11]、单只 ETF^[12]、ETF 组合^[13]。本文认为成份股完全复制不具备可操作性, 部分复制存在交易成本和跟踪误差的权衡, 同时容易受到人为选股失误的影响。ETF 虽然具有交易成本低的优势, 但是考虑到流动性因素, ETF 投资组合应是期现套利中现货替代工具的最好选择。

本文 ETF 组合各 ETF 的权重确定方法是基于 Beasley 等^[14] 所使用的修正跟踪误差指标来判断。令 $R_{i,t}$ 为指数的第 t 分钟对数收益率, $R_{i,t}$ 为 ETF 组合中第 i 只 ETF 第 t 分钟的对数收益率, 则在第 t 分钟时的收益率偏差 $TE_t = R_{i,t} - \sum_{i=1}^n w_i R_{i,t}$, 其中 w_i 为第 i 只 ETF 在 ETF 组合中的权重, ETF 组合的分钟均跟踪误差 $ATE = \frac{1}{T} \sqrt{\sum_{t=1}^T [R_{i,t} - \sum_{i=1}^n w_i R_{i,t}]^2}$ 。

收稿日期: 2012-09-21; 修回日期: 2012-11-02。

作者简介: 王飞(1965 -) 男, 江苏南通人, 副教授, 主要研究方向: 高频交易算法、证券投资分析; 孙维尧(1988 -) 男, 吉林白山人, 硕士研究生, 主要研究方向: 高频交易算法、证券投资分析。

ETF 组合应满足以下最优化目标函数:

$$\text{Minimize } ATE = \frac{1}{T} \sqrt{\sum_{i=1}^T \left[R_{i,t} - \sum_{j=1}^n w_j R_{i,j} \right]^2}; \quad \sum_{i=1}^n w_i = 1$$

1.2 基于 B-S 方程的股指期货期现套利模型

股指期货的理论价值 f 等于连续复利条件下股指现货价值 S 的未来值, 表示为:

$$f = S e^{(r-q)t} \quad (1)$$

其中: r 为无风险利率, q 为指数股息率, t 为距合约到期日的时点, $e^{(r-q)t}$ 为连续复利因子。

构建一个包含股指期货空头和 ETF 多头的投资组合。为了表达直观, 在以下推导中令 ETF 为 ETF 的价格序列。投资组合的初始价值 $\Pi = 0$:

$$\Pi = -f + \alpha ETF \quad (2)$$

其中 α 为使投资组合无风险条件下 ETF 的投资比例, 其具体形式由下文给出。

连续动态条件下:

$$d\Pi = -df + \alpha dETF \quad (3)$$

式 (1) 为基于标的资产 S (即股指现货) 的衍生产品理论定价公式, 由伊藤引理得:

$$df = \left(\frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S} \sigma S dZ \quad (4)$$

其中: μ 为 S 的期望, σ^2 为 S 的方差; dZ 是一个维纳过程, 是 f 基于 S 风险溢价的表示。

同时, 在连续动态条件下, ETF 服从以下过程:

$$dETF = \mu_{ETF} ETF dt + \sigma_{ETF} ETF dZ \quad (5)$$

其中: μ_{ETF} 为 ETF 的期望, σ_{ETF} 为 ETF 的标准差; dZ 为一个维纳过程, 是 ETF 基于 S 风险溢价的表示。因为 dZ 的标的资产相同, 所以式 (4)、(5) 的 dZ 维纳过程应是一样的。

在投资组合无风险条件下, 动态的 $d\Pi$ 中股指期货部分 ETF 中含 dZ 的项应该相互抵消, 因此:

$$\alpha = \frac{\frac{\partial f}{\partial S} \sigma_{ETF} ETF}{\frac{\partial f}{\partial S} \sigma S} \quad (6)$$

将式 (3) ~ (6) 联立, 得无风险条件下的投资组合动态价值:

$$d\Pi = - \left(\frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \alpha \sigma_{ETF} ETF dt \quad (7)$$

无套利条件下投资组合的收益率应该为无风险收益率 r , 即:

$$d\Pi = r\Pi dt \quad (8)$$

将式 (2)、(7) 代入式 (8), 整理得:

$$rf - \frac{\partial f}{\partial S} \mu S - \frac{\partial f}{\partial t} - \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 = \alpha(r - \mu_{ETF}) ETF \quad (9)$$

因为 $f = S e^{(r-q)t}$, 对 f 求偏导数, 则式 (9) 可变为:

$$rS e^{(r-q)t} - \mu S e^{(r-q)t} - (r - q) S e^{(r-q)t} = \alpha(r - \mu_{ETF}) ETF \quad (10)$$

式 (10) 是式 (8) 的具体形式, 等式中的左边部分是为保证无套利机会下投资组合中关于股指期货的部分, 等式右边是投资组合中关于 ETF 的部分。该无套利等式只有在 $t = dt$ 时成立, 因此该投资组合的无套利情况只有在极短的时间内成立, 一般情况下, 式 (10) 可写成:

$$rS e^{(r-q)t} - \mu S e^{(r-q)t} - (r - q) S e^{(r-q)t} = c + \alpha(r - \mu_{ETF}) ETF + \varepsilon \quad (11)$$

在式 (11) 中, $S e^{(r-q)t}$ 为股指期货的理论价值, 令 F 为股指期货的市价序列, 取代 $S e^{(r-q)t}$ 项, 整理, 则有以下等式:

$$(q - r) F = \alpha(r - \mu_{ETF}) ETF \quad (12)$$

令 $\beta = \alpha \frac{r - \mu_{ETF}}{q - r}$, 则式 (12) 可简化为 $F = \beta ETF$, 令基于实

际市场数据的投资组合在第 t 分钟的实际价差为 $spread_t$, 则:

$$spread_t = F_t - \beta ETF_t \quad (13)$$

1.3 算法交易系统

期现套利模型需要部署在算法交易系统中, 算法交易系统为算法提供输入参数、接受算法输出。一般套利算法交易系统的输入参数包括交易指令、历史行情、实时行情和成交回报。其中交易指令需要人工制定, 实时行情需要通过本地系统按照各交易所行情接口接收下来, 并以 DBF 格式存到行情数据库中, 其中现货每 3 s 更新一次, 期货最高频率为 500 ms 一次。

本文的交易算法系统要求算法识别两类套利情景: $spread_t > \mu_{Spread}$ 和 $spread_t < \mu_{Spread}$ (其中 t 为步长, $1 < t < T$; μ_{Spread} 为某期间内的 $spread_t$ 均值), 前者为正向套利状态, 后者为反向套利状态; 算法交易指令涉及三类六个参数, 分别为 λ_i 、 χ_i 和 δ_i ($i = 1, 2$)。其中 λ_i 和 χ_i 用以确定无套利区间上下边界, χ_1 和 χ_2 用以确定触发止损指令的价差, δ_1 和 δ_2 用以确定触发获利平仓的价差。算法需要的其他参数如下: 持仓状态 $open$, $open = 1$ 代表持仓, 0 代表空仓。算法步骤如下:

- 1) 人工输入参数 r 、 q , 根据特定期间历史数据计算 β 、 $spread$ 的历史序列, 计算 $spread$ 的均值 μ_{Spread} 和方差 σ_{spread} 。
- 2) 若 $t = T$, $open = 1$ 强制平仓 $open \leftarrow 0$; 若 $t < T$ 进入 3)。
- 3) 根据式 (13) 计算 $spread$, $open = 0$ 时识别建仓指令, $open = 1$ 时识别获利平仓和止损指令, $t \leftarrow t + 1$ 转 2)。

具体交易指令如下:

1) 建仓指令:

当 $\mu_{Spread} + \lambda_1 \sigma_{spread} < spread_t \leq \mu_{Spread} + \chi_1 \sigma_{spread}$ 时, 卖空 1 单位股指期货同时将 $300 \times \beta ETF_t$ 的资金按 w_i 投资各 ETF, $open \leftarrow 1$;

当 $\mu_{Spread} - \chi_2 \sigma_{spread} \leq spread_t < \mu_{Spread} - \lambda_2 \sigma_{spread}$ 时, 买入 1 单位股指期货同时将 $300 \times \beta ETF_t$ 的资金按 w_i 融券做空 ETF, $open \leftarrow 1$ 。

2) 获利平仓指令:

正向状态开仓后, 若 $spread_t < \mu_{Spread} - \delta_1 \sigma_{spread}$, 平仓止盈, $open \leftarrow 0$ 。

反向状态开仓后, 若 $spread_t > \mu_{Spread} + \delta_2 \sigma_{spread}$, 平仓止盈, $open \leftarrow 0$ 。

3) 止损指令: 若 $spread_t > \mu_{Spread} + \chi_1 \sigma_{spread}$ 或 $spread_t < \mu_{Spread} - \chi_2 \sigma_{spread}$, 平仓止损 $open \leftarrow 0$ 。

2 实证检验

本文利用 Excel 读取行情数据并作为算法交易系统的运算平台, 实践中亦可将 DBF 数据转发至 Matlab 或 SAS。本文采用 2012 年 8 月 17 日一分钟行情数据, SHIBOR 当日的 1M 年化值为 0.040 517, 沪深 300 指数股息率为 0.020 8, 股指期货合约为 IF1209。

2.1 ETF 投资组合的构建

本文选择了三只流动性较好的 ETF: 50ETF、180ETF 和 300ETF。根据上文 ETF 投资组合最优化目标函数, 本文使用广义既约梯度法 (Generalized Reduced Gradient, GRG) 进行求解, 输出结果如表 1。

表 1 ETF 投资组合各 ETF 的权重

ETF 品种	50ETF	180ETF	300ETF
权重	0.147 794	0.333 333	0.518 871

2.2 期现套利模型实证结果

图 1 为 $spread_t$ 序列的计算结果图形。

本文在确定交易指令参数时假设价差上下波动风险是对称的,如图 1,当 $spread$ 波动超过一个标准差时执行建仓指令,即 $\lambda_1 = \lambda_2 = 1$; 当 $spread$ 波动超过三个标准差时执行止损指令,即 $\chi_1 = \chi_2 = 3$; 建仓后当 $spread_t$ 回到均值时执行平仓指令,即 $\delta_1 = \delta_2 = 0$ 。根据这三类预订交易指令,本文模型在 2012 年 8 月 17 日的交易结果统计如表 2 所示。

如表 2 显示,当日该算法交易系统共执行了 57 个建仓指令,可见基于 Black-Scholes 方程的股指期货期现套利模型能够实现日内的高频交易,交易频率较一般套利模型有很大提高,同时仅有 5.26% 的止损发生率,说明该模型在提高交易频率的同时避免了相当数量的伪信号,具有很高的交易信号识别能力。本文所设定的交易指令可以将收益和损失锁定在具体的范围内,如 8 月 15 日内的单笔收益率为万分之三,单笔损失约为万分之七,可见本模型同常规的统计套利模型一

样,在市场稳定的交易日,止损发生率较低,可以实现多次的少量收益进而确保整体收益,而在市场波动较大的交易日,止损发生率会显著提高导致出现巨大亏损。因此需要投资经理根据现实的市场状况考虑是否利用该模型进行交易,而获利的关键并不在于单笔收益率的大小,而是在于发掘更多的交易信号同时控制止损率。

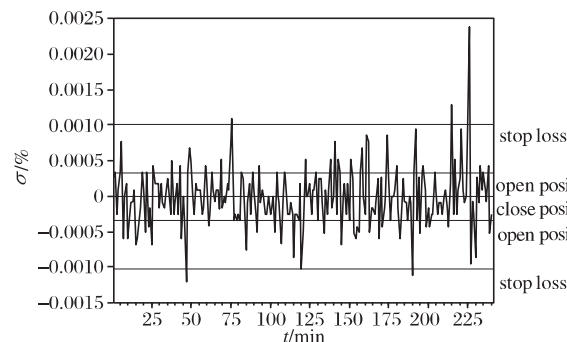


图 1 套利示意图

表 2 套利结果统计表

套利方向	建仓次数	止损数	单笔收益率/%	单笔损失率/%	单日净收益率/%	年化收益率/%	止损发生率/%
正向套利	30	1	0.033	0.067	1.523	2051.02	5.263
反向套利	27	3					

3 结语

本文所构建的基于 Black-Scholes 方程的股指期货期现套利模型及交易算法,在市场相对稳定的前提下,能够很好地捕捉交易机会实现客观的整体收益,但是如果市场波动较大,本文模型的适用性就会受到限制。所以投资经理在利用本文模型进行交易前应先对市场状况进行分析以判断是否符合模型的获利条件。

参考文献:

- [1] CORNELL B, FRENCH K R. The pricing of stock index futures [J]. Journal of Futures Markets, 1983, 3(1): 1-14.
- [2] CHU Q C, HSIEH W G, TSE Y. Price discovery on the S&P 500 index markets: an analysis of spot index, index futures, and SPDRS [J]. International Review of Financial Analysis, 1999, 8(1): 21-34.
- [3] RICHIE N, DAIGLER R, GLEASON K. The limits to stock index arbitrage: examining S&P 500 futures and SPDRS [J]. Journal of Futures Market, 2008, 28(12): 1182-1205.
- [4] ELLIOTT R, van der HOEK J, MALCOLM W P. Pairs trading [J]. Quantitative Finance, 2005, 5(3): 271-276.
- [5] LIN Y, MICHAEL M, CHANDRA G. Loss protection in pair trading through minimum profit bounds: a cointegration approach [J]. Journal of Applied Mathematics and Decision Sciences, 2006, 2006: 1-14.
- [6] 仇中群,程希骏. 基于协整的股指期货跨期套利策略模型[J]. 系统工程, 2008, 26(12): 26-29.
- [7] 张连华. 基于高频数据的股指期货期现统计套利程序交易[J]. 计算机应用软件, 2011, 28(9): 93-95.
- [8] 刘华. 基于统计套利的 ETF 期限套利方法应用研究[D]. 大连: 大连理工大学, 2008.
- [9] TRAPLETTI A, GEYER A, LEISCH F. Cointegration and exchange market efficiency: an analysis of high frequency data [EB/OL]. (1999-08) [2012-09-01]. <http://www.ci.tuwien.ac.at/~leisch/docs/papers/wp5>.
- [10] 雷泽,秦昌贵. 股指期货: 套利空间大利润高[J]. 资本市场, 2007(5): 68-69.
- [11] 马理,卢烨婷. 沪深 300 股指期货期现套利的可行性研究[J]. 财贸研究, 2011, 88(1): 89-93.
- [12] 崔健军. 上证 50ETF 与沪深 300 期指套利[N]. 期货日报, 2007-07-03(2).
- [13] 黄晓坤,侯金鸣. 利用 ETF 类基金进行股指期货套利的方法研究[J]. 统计与决策, 2009, 132(18): 132-134.
- [14] BEASLEY J E, MEADE N, CHANG T-J. An evolutionary heuristic for the index tracking problem[J]. European Journal of Operational Research, 2003, 148(3): 621-643.

(上接第 320 页)

- [6] U. S. Department of Transportation, Bureau of Transportation Statistics. Transportation statistics annual report 2007 [R]. Washington, DC: U. S. Department of Transportation, Bureau of Transportation Statistics, 2007.
- [7] U. S. Department of Transportation, Bureau of Transportation Statistics. National transportation statistics 2007 [M]. Washington, DC: U. S. Department of Transportation, Bureau of Transportation Statistics, 2007.
- [8] LIU W, KING H. RFID antenna design of highway ETC in ITS[C]// Proceedings of 2006 7th International Symposium on Antennas, Propagation and EM Theory. Guilin: [s. n.], 2006: 300-303.
- [9] U. S. Department of Transportation, Bureau of Transportation Statistics. National transportation statistics 2011 [R]. Washington, DC: U. S. Department of Transportation, Bureau of Transportation Statistics, 2011.
- [10] U. S. Department of Transportation, Bureau of Transportation Statistics. National transportation statistics 2012 [R]. Washington, DC: U. S. Department of Transportation, Bureau of Transportation Statistics, 2012.
- [11] 陈荫三,李彬,肖润谋. 2010 年中国高速公路网运输状态[J]. 交通运输工程学报, 2011, 6(11): 68-73.
- [12] 冯丽云,洪玫,杨秋辉,等. 基于逆向查询处理算法的数据库系统测试数据生成方法[J]. 计算机应用, 2011, 31(4): 948-951.