

基于随机价差法的配对交易研究

蔡 燕¹,王 林²,许莉莉¹

(1.西安交通大学,陕西 西安 710049;2.山西证券研究所,山西 太原 030000)

摘要:配对交易是一种风险中性的均值回复投资策略。当配对的品种价差偏离历史均值时,则做空价格较高的品种同时买进价格较低品种,等待它们回归到长期均衡关系,平仓。目前国内主要使用协整理论研究配对机会,协整方法和随机价差法都是均值回复方法。在协整的基础上引入两种随机价差模型(价差服从O-U随机过程和基于Elliott框架)研究沪深300股指期货和上证180ETF的配对机会并给出了预测曲线,结果显示价差服从O-U过程的模型预测结果与真实数据的相关性高于基于Elliott框架的预测结果,并且标准差与真实数据的标准差更接近。

关键词:期货交易;配对交易;均值回复;协整;股指期货;ETF

文章编号:1003-4625(2012)08-0030-06 **中图分类号:**F830.9 **文献标志码:**A

一、引言

2010年4月16日,中国金融期货交易所推出沪深300股指期货合约,该合约以沪深300指数为标的,沪深300指数是根据流动性和市值规模从沪深两市中选取300只A股股票作为成分股编制的。股指期货能增加资本市场流动性;对现货市场具有价格发现功能;同时能规避市场的系统性风险,所以股指期货一经推出便立刻成为金融市场交易最活跃的品种。

ETF基金,即交易型开放式指数基金,是一种在交易所上市交易的、基金份额可变的开放式基金,其主要目的是跟踪标的指数的变化。投资者既可以在二级市场买卖ETF份额,又可以向基金管理公司申购或赎回ETF份额,不过申购赎回必须以一篮子股票(或有少量现金)换取基金份额或者以基金份额换回一篮子股票(或有少量现金)。由于同时存在二级市场交易和申购赎回机制,投资者可以在ETF二级市场交易价格与基金单位净值之间存在差价时进行套利交易。套利机制的存在,可使ETF避免封闭式基金普遍存在的折价问题。A股市场上交易活跃的ETF有上证50ETF、上证180ETF、深圳100ETF以及刚刚推出的沪深300ETF。沪深300ETF是完全复制沪深300指数的,在二级市场具备实时交易功能,这就为期限套利提供了更为方便的工具。

配对交易是一种统计套利方法。统计套利是指将套利建立在对历史数据进行统计分析的基础上,估计相关变量的概率分布,并结合基本面数据进行分析以用来指导套利交易。配对交易是指从市场上找出历史价格走势相近的品种进行配对,当配对的品种价格差偏离历史均值时,则做空股价较高的股票同时买进股价较低的股票,等待它们回归到长期均衡关系,由此赚取两股票价格收敛的报酬。

国内许多学者用协整理论研究套利交易以及研究股指期货与ETF间的统计套利机会。例如张敏、徐坚^[1]研究了利用ETF构建沪深300指数期货的现货组合的可能性,对几种ETF组合与沪深300指数的相关性以及它们跟踪沪深300指数的跟踪误差进行了分析,结果显示,使用ETF组合作为现货组合来模拟沪深300指数能取得良好效果。薛战忠^[2]通过对沪深300股指期货指数和上证50ETF的历史数据进行分析,发现两者都是非平稳时间序列数据,因而利用协整检验,发现两者具有长期均衡的协整关系,为套利交易提供理论基础。通过历史数据找出两者的回归关系,找出套利机会出现的条件和套利手段,为投资者投资提供建议。国外也有许多学者研究配对交易。Gatev等^[3]提出了使用与给定股票的标准化序列之间的偏差平方和最小的股票构成股票对,相应的门限值是配对股票标准化价格的差的历史标准

收稿日期:2012-06-04

作者简介:蔡燕(1987-),女,四川自贡人,硕士研究生,研究方向:金融数学;王林(1986-),男,重庆人,硕士,研究员,研究方向:金融工程;许莉莉(1988-),女,山东菏泽人,硕士研究生,研究方向:金融数学。

差的两倍。Frazzo等^[4]认为配对股票的选择应将股票按长、短期的相关系数来进行同行业内的股票配对,再以两股票现在价格相较与其长期平均数而言的标准差来作为交易评判的标准。Vidyamurthy^[5]认为在执行配对交易策略时,应先就股票基本面或历史数据挑选出潜在具有协整关系的配对股票,并以协整系数及均值来构建两股票价格的线性关系。Elliott等^[6]给出了运用随机价差模型进行配对交易的一个基本的分析框架。Betram^[7]在假设对数价差满足零均值的Ornstein-Uhlenbeck过程下,对平仓时间和期望持仓时间做了分析。Cummins^[8]运用爱尔兰交易所的数据对Betram提出的方法做了实证分析,Cummins发现假设对数价差过程满足的O-U过程会产生较大的误差,这种误差体现在该假设可能出现对单位时间期望收益的过高估计和对交易策略持续时间的过低估计。

目前,国内研究配对交易通常指的是用计量经济学的方法分析套利机会,如协整理论,本文采用一种新的思路(随机价差法)寻找配对机会。随机价差法与协整理论具有相似之处(都可认为是均值回复方法或者逆势交易方法)。本文对实验数据也做了协整检验,结果显示具有协整关系,也即是说价差具有均值回复特性,文章创新之处是利用价差具有均值回复特性对价差做进一步假定,即价差服从某均值回复的随机过程。同时,本文的另一目的是:为投资者参与沪深300ETF与股指期货套利提供一种理论方法。本文中,实证数据的现货标的是上证180ETF,可以预期刚刚推出的沪深300ETF与股指期货的相关性强于上证180ETF与股指期货,故配对交易效果更好。

第一章是绪论,主要介绍沪深300股指期货、ETF、配对交易概念和国内外研究现状以及研究意义;第二章讲述随机价差理论方法,(一)节介绍价差服从O-U过程的均值回复理论,(二)节介绍基于Elliott框架的均值回复理论;第三章为实证研究,检验数据的协整关系,应用样本数据对模型进行参数估计,并根据估计的参数值进行测试数据的预测;第四章是结论与展望。

二、理论方法介绍

传统价差套利或价比套利依据当前价差与历史价差偏差,主要根据套利者的经验和直觉。当前研究热点协整套利的基本思想是对价格序列做协整检验,若判定价格序列间存在长期协整关系,则根据此协整关系套利,即若价格偏离均衡值,开仓;等待价格回归均衡值,平仓。协整套利本质上利用了价差

具有均值回复特性。均值回复过程是一类特殊的随机过程,该随机过程以回复速度靠拢长期均值水平。本文采用的随机价差法本质上也是利用价差的均值回复特性。随机价差法假定价差服从某均值回复过程。第二章(一)节假定价差服从Ornstein-Uhlenbeck过程^[9],使用最大似然函数估计O-U过程的参数,并通过蒙特卡罗模拟获得估计参数的均值。第二章(二)节介绍Elliott等人提出的基本框架,首先,使用Shumway and Stoffer方法估计随机价差模型的参数;然后,使用Elliott和Krishnamurthy滤波方法估计随机价差模型的参数。

(一)价差服从O-U过程的均值回复理论

假定价差观测值 $\{y_k|k=0,1,2,\dots\}$ 服从如下O-U过程:

$$dy_t = k(\theta - y_t)dt + \sigma dw_t \quad (1)$$

其中: y_t 表示t时刻的价差; dy_t 表示t时刻价差的变化; k 和 θ 分别表示价差的均值回复速度和长期均值回复水平, σ 表示价差的波动大小; dw_t 表示维纳增量过程。 $\Theta=(k,\theta,\sigma)$ 为待估计参数。

已知s时刻的价差 y_s ,则t时刻价差 y_t 的条件期望为 $E(y_t|y_s) = \theta + e^{-k(t-s)}(y_s - \theta)$,可知相应方差为 $D(y_t|y_s) = \frac{\sigma^2}{2k}(1 - e^{-2k(t-s)})$ 条件概率密度函数为:

$$P(y_t|y_s) = \frac{k}{\sqrt{\pi\sigma^2(1 - e^{-2k(t-s)})}} \exp\left(\frac{-k[(y_t - \theta) - (y_s - \theta)e^{-k(t-s)}]^2}{\sigma^2(1 - e^{-2k(t-s)})}\right) \quad (2)$$

令 $L = \prod_{i=1}^{N-1} P(y_i|y_{i-1})$,N表示样本个数,则得到对数似然函数:

$$\ln L = \frac{N-1}{2} \ln\left(\frac{\alpha}{\pi}\right) + \sum_{i=1}^{N-1} -\alpha [(y_i - y_{i-1}) + \beta (y_{i-1} - \theta)]^2 \quad (3)$$

$$\text{其中: } \alpha = \frac{k}{\sigma^2(1 - e^{-2k\Delta t})}, \beta = 1 - e^{-k\Delta t}, t-s = \Delta t = 1$$

从而可以估计出 $\hat{\Theta} = (\hat{k}, \hat{\theta}, \hat{\sigma}) = \arg \max_{\Theta} \ln L$ (Θ)。

为了进一步保证参数估计的准确性,我们使用蒙特卡罗模拟方法。蒙特卡罗模拟是一种通过设定随机过程,反复生成时间序列,计算参数估计量和统计量,进而研究其分布特征的方法。具体地,当系统中各个单元的可靠性特征量已知,但系统的可靠性过于复杂,难以建立可靠性预计的精确数学模型或模型太复杂而不便应用时,可用随机模拟法近似计算出系统可靠性的预计值;随着模拟次数的增多,其预计精度也逐渐增高。该方法以概率和统计理论方

法为基础的一种计算方法,是使用随机数来解决很多计算问题的方法。从蒙特卡罗方法的思路可以看出,该方法回避了结构可靠度分析中的数学困难,只要模拟的次数足够多,就可以得到一个较精确的解。将模型离散化:

$$x_{t+1}=x_t+\hat{k}(\hat{\theta}-x_t)\Delta t+\hat{\sigma}\sqrt{\Delta t}w_{t+1} \quad (4)$$

$$\sqrt{\Delta t}w_{t+1}\rightarrow N(0,\Delta t)$$

本文模拟多条样本路径,再运用每条样本路径分别进行参数估计,最后求取平均值得到估计参数的均值:

$$\tilde{\Theta}=(\tilde{k},\tilde{\theta},\tilde{\sigma})=\frac{1}{N}\sum_{i=1}^N\hat{\Theta}_i \quad (5)$$

(二)基于Elliott配对交易框架的均值回复理论

Elliott提出的配对交易框架与(一)节的配对交易理论有明显不同。在(一)节中,直接假定观测数据服从O-U过程,而Elliott假定观测数据是真实数据的随机扰动,真实数据服从O-U过程,即真实价差 $\{x_k|k=0,1,2,\dots\}$ 服从等式, $\{x_k|k=0,1,2,\dots\}$ 的离散形式如下:

$$x_{t+1}-x_t=k(\theta-x_t)\Delta t+\sigma\sqrt{\Delta t}w_{t+1} \quad (6)$$

观测值序列 $\{y_k|k=0,1,2,\dots\}$ 为真实过程的扰动,如下:

$$y_k=x_k+Dw_k \quad (7)$$

等式(1)等价于:

$$x_{k+1}=A+Bx_k+C\epsilon_{k+1} \quad (8)$$

其中: $A=\kappa\theta\Delta t$, $B=1-\kappa\Delta t$, $C=\sigma\sqrt{\Delta t}$, $0<B<1$

(7)——(8)是Elliott等人提出的随机价差法的基本模型。Elliott等人使用卡尔曼滤波方法产生真实数据(不同于观测数据)。基本步骤如下:给定估计参数初值 $\Omega_0=(A,B,C,D)$ 和初始值 \hat{x}_0 , $\Sigma_{0|0}$,计算等式、得到 (\hat{x}_k,R_k) ,从而根据Elliott的卡尔曼滤波迭代过程计算出 (\hat{x}_{k+1},R_{k+1}) 。

$$\mu_k=\hat{x}_k\equiv\hat{x}_{k|k}=E[x_k|\psi_k] \quad (9)$$

$$R_k=\sum_{k|k}\equiv E[(x_k-\hat{x}_k)^2|\psi_k] \quad (10)$$

$$\hat{x}_{k+1|k}=A+B\hat{x}_k \quad (11)$$

$$\sum_{k+1|k}=B^2\sum_{k|k}+C^2 \quad (12)$$

$$K_{k+1}=\frac{\sum_{k+1|k}}{(\sum_{k+1|k}+D^2)} \quad (13)$$

$$\hat{x}_{k+1}=\hat{x}_{k+1|k+1}=\hat{x}_{k+1|k}+K_{k+1}(y_{k+1}-\hat{x}_{k+1|k}) \quad (14)$$

$$R_{k+1}=\sum_{k+1|k+1}=D^2K_{k+1}=(1-K_{k+1})\sum_{k+1|k} \quad (15)$$

由卡尔曼滤波产生真实数据之后,我们需要获得模型(7)——(8)的估计参数 $\Omega=(A,B,C,D)$ 。使用

EM算法估计模型的参数,EM算法是迭代算法,如果 $\hat{\Omega}_0$ 是初始估计值,通过(16)——(17)两步计算 $\hat{\Omega}_j,j=1,2,\dots$ 。

$$Q(\Omega,\hat{\Omega})=E_{\hat{\Omega}}[\ln\frac{dP_{\Omega}}{dP_{\hat{\Omega}}}|\psi_N] \quad (16)$$

$$\Omega_{j+1}\in\arg\max_{\Omega}Q(\Omega,\hat{\Omega}_j) \quad (17)$$

参数估计的方法有很多,例如:极大似然估计、矩估计、贝叶斯估计^[10]等。本文这两种方法都是基于EM算法的,在此只做简要介绍^[11]。下面第1小节介绍Shumway and Stoffer光滑方法^[12];第2小节介绍Elliott和Krishnamurthy滤波方法^[13]。

1.Shumway and Stoffer方法

定义如下记号:

$$\hat{x}_{k|N}=E[x_k|\psi_N] \quad (18)$$

$$\sum_{k|N}=E[(x_k-\hat{x}_{k|N})^2|\psi_N]=E[(x_k-\hat{x}_{k|N})^2] \quad (19)$$

$$\sum_{k-1,k|N}=E[(x_k-\hat{x}_{k|N})(x_{k-1}-\hat{x}_{k-1|N})] \quad (20)$$

Shumway和Stoffer算法的具体实现如下:

$$\vartheta_k=\frac{B\sum_{k|N}}{\sum_{k+1|k}} \quad (21)$$

$$\hat{x}_{k|N}=\hat{x}_{k|k}+\vartheta_k[\hat{x}_{k+1|N}-(A+B\hat{x}_{k|N})] \quad (22)$$

$$\sum_{k|N}=\sum_{k|k}+\vartheta_k^2[\sum_{k+1|N}-\sum_{k+1|k}] \quad (23)$$

$$\sum_{k-1,k|N}=\vartheta_{k-1}\sum_{k|k}+\vartheta_{k-1}\vartheta_k[\sum_{k,k+1|N}-B\sum_{k|k}] \quad (24)$$

$$\sum_{N-1,N|N}=B(1-K_N)\sum_{N-1|N-1} \quad (25)$$

该方法是向后迭代算法,初始迭代值 $\hat{x}_{N|N}$ 和 $\sum_{N|N}$ 可由卡尔曼滤波获得。给定 $\Omega_j=(A,B,C^2,D^2)$ 和卡尔曼滤波的初值 $\hat{x}_0=j^{-1}\hat{x}_{0|N}$ 、 $\sum_{0|0}=j^{-1}\sum_{0|N}$ 。求解 $\hat{\Omega}_{j+1}=(\hat{A},\hat{B},\hat{C}^2,\hat{D}^2)$ 的过程如下:

$$\hat{A}=\frac{\alpha\gamma-\delta\beta}{N\alpha-\delta^2} \quad (26)$$

$$\hat{B}=\frac{N\beta-\gamma\delta}{N\alpha-\delta^2} \quad (27)$$

$$\hat{C}^2=\frac{1}{N}\sum_{k=1}^N[(x_k-\hat{A}-\hat{B}x_{k-1})^2|\psi_N] \quad (28)$$

$$\hat{D}^2=\frac{1}{N+1}\sum_{k=0}^N[(y_k-x_k)^2|\psi_N] \quad (29)$$

其中:

$$\alpha=\sum_{k=1}^NE[x_k^2|\psi_N]=\sum_{k=1}^NE[\sum_{k-1|N}+x_{k-1|N}^2] \quad (30)$$

$$\beta=\sum_{k=1}^NE[x_kx_{k-1}|\psi_N]=\sum_{k=1}^NE[\sum_{k-1,k|N}+x_{k-1|N}x_{k|N}] \quad (31)$$

$$\gamma=\sum_{k=1}^N\hat{x}_{k|N} \quad (32)$$

$$\hat{\delta} = \sum_{k=1}^N \hat{x}_{k-1|N} = \gamma - \hat{x}_{N|N} + \hat{x}_{0|N} \quad (33)$$

2.Elliott和Krishnamurthy滤波方法

同第1小节一样,我们以 $\Omega_j=(A,B,C^2,D^2)$ 为初始值,求解 $\hat{\Omega}_{j+1}=(\hat{A},\hat{B},\hat{C}^2,\hat{D}^2)$ 。该算法在计算效率上比Shumway和Stoffer算法好。求解过程如下:

$$H_k^0 = \sum_{l=0}^k x_l^2, \hat{H}_k^0 = E[H_k^0 | \psi_k]$$

$$H_k^1 = \sum_{l=1}^k x_l x_{l-1}, \hat{H}_k^1 = E[H_k^1 | \psi_k]$$

$$H_k^2 = \sum_{l=0}^k x_{l-1}^2, \hat{H}_k^2 = E[H_k^2 | \psi_k]$$

$$J_k = \sum_{l=0}^k x_l y_l, \hat{J}_k = E[J_k | \psi_k]$$

$$I_k^0 = \sum_{l=0}^k x_l, \hat{I}_k^0 = E[I_k^0 | \psi_k]$$

$$I_k^1 = \sum_{l=1}^k x_{l-1}, \hat{I}_k^1 = E[I_k^1 | \psi_k]$$

$$Y_k = \sum_{l=0}^k y_l^2$$

如果 $E = E_{\hat{\Omega}_j}$,则有:

$$\hat{A} = [1 - \frac{(\hat{I}_N^1)^2}{\hat{H}_N^2}]^{-1} [\hat{I}_N^0 - \frac{\hat{H}_N^1 \hat{I}_N^1}{\hat{H}_N^2}] \quad (34)$$

$$\hat{B} = \frac{1}{\hat{H}_N^2} [\hat{H}_N^1 - \hat{A} \hat{I}_N^1] \quad (35)$$

$$\hat{C}^2 = \frac{1}{T} [\hat{H}_N^0 + T \hat{A}^2 + \hat{H}_N^2 \hat{B}^2 - 2 \hat{A} \hat{I}_N^0 + 2 \hat{A} \hat{B} \hat{I}_N^1 - 2 \hat{B} \hat{H}_N^1] \quad (36)$$

$$\hat{D}^2 = \frac{1}{T+1} [Y_N - 2 \hat{J}_N + \hat{H}_N^0] \quad (37)$$

三、实证研究

(一)数据说明

本文选取沪深300股指期货当月连续合约价格和上证180ETF作为研究对象。选取自2010年4月16日股指期货推出之日起,到2012年4月9日的日收盘数据作为样本总体,各计479个。其中将2010年4月16日到2011年12月2日之间的400个数据作为实验样本,2011年12月2日以后的79个数据作为测试样本。配对数据满足同时发生的要求。

(二)数据处理

1.相关性分析

计算股指期货连续合约和上证180ETF的日收

盘价的单手价格。图1是两价格的历史走势^①。它们的变动趋势十分相似,相关系数为0.995。

二、协整检验

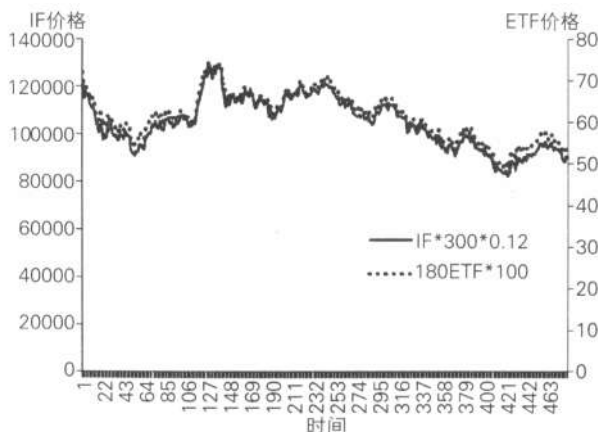


图1 股指期货连续合约和上证180ETF的日收盘价的单手价格趋势图

许多时间序列属于非平稳序列,其变动性会随着时间的变化而变化。格兰杰在20世纪80年代提出的协整理论发现,把两个或两个以上非平稳的时间序列进行特殊组合后可能呈现出平稳性。协整理论的主要研究对象是在两个(或多个)非平稳时间序列中寻找一种均衡关系,该理论的提出对于用非平稳经济变量建立计量经济模型,以及检验这些变量之间的长期均衡关系具有非常重要的意义,而且其应用也远远超出了对线性回归的诊断。在许多情况下,经济理论告诉我们两个变量应该是协整的,对协整性的检验也就是对经济理论正确与否的检验。

协整的应用意义在于反映经济变量之间的长期均衡关系。例如若两组时间序列数据满足协整关系,那么从长期上看这两组数据存在均衡关系,任何对均衡位置的偏离都会促使其向均衡位置回归。基于此方法的统计套利称为协整套利。

首先检验数据的平稳性,观察变量的时序图,初步判断模型是否需要常数项和趋势项,最后选定模型进行对原始序列做ADF单根检验。表1是检验结果,从表中可以知道:原始数据均为一阶单整,故非平稳。采用E-G两步法^[14]检验数据的协整关系:

对原始数据做线性回归: $Y_t = \alpha + \beta X_t + u_t$ 。

其中: Y 表示12%的保证金时一手沪深300股指期货合约的价格序列; X 表示一手上证180ETF的价格序列; u_t 服从标准正态分布。

价差^②序列 $\hat{u}_t = Y_t - (\hat{\alpha} + \hat{\beta} X_t)$,检验价差 \hat{u}_t 的平

①注:图1是这样计算的:股指期货一点代表300元,假设12%保证金;上证180ETF一手100份。

② \hat{u}_t 通常被叫做残差,为了与前文叫法保持一致性,且不失一般性,把 \hat{u}_t 叫做价差。

稳性。表2显示 \hat{u}_t 是平稳的。因此沪深300股指期货与上证180ETF价格序列存在协整关系,也即是说它们存在长期均值回复关系,本文不仅从计量经济学的理论做套利研究,而且假定价差序列 \hat{u}_t 服从第二章提出的两个均值回复模型,第(三)节分别对这两个模型做了实证研究。

表1 原始序列平稳性检验

	180ETF			IF		
	Statistic	Prob.	平稳性	Statistic	Prob.	平稳性
原始序列	1.53832	0.4634	No	1.21356	0.5451	No
一阶差分序列	157.532	0.0000	Yes	157.559	0.0000	Yes

表2 价差序列平稳性检验

Statistic	Prob.	1%level	5% level	10% level	平稳性
-4.031056	0.0014	-3.446567	-2.868583	-2.570588	Yes

(三)数值模拟结果

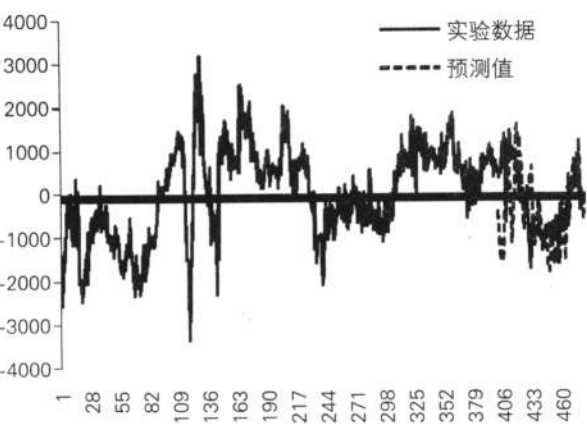


图2 价差服从O-U过程的预测图

运用极大似然估计对第二章第(一)节中的模型进行参数估计,基于样本数据的参数估计结果为 $\hat{\Theta}=(0.169, 72.96, 620.3)$ (等价于 $A=12.2, B=0.83, C=620.3$);蒙特卡罗模拟得最终结果为 $\tilde{\Theta}=(0.196, 72.27, 679.85)$ (等价于 $A=14.16, B=0.8, C=679.85$);在对第(二)节中模型-的参数估计时,Shumway和Stoffer方法的参数估计结果为 $A=7.2, B=0.84, C=296.6, D=600.8$;Elliott和Krishnamurthy方法的参数估计结果为 $A=19.8, B=0.89, C=210.9, D=190.9$ 。利用上述参数估计值,我们对测试数据(2011年12月2日到2012年4月16日的79个数据)进行预测。图2-图4分别给出了基于上述三类参数估计的一次预测结果。可以看出预测值与观测数据变化趋势有正相关性。图5给出Elliott和Krishnamurthy方法中参数B收敛到0.89,若B介于(0,1),价差序列是收敛

的,否则发散,因此Elliott和Krishnamurthy方法是收敛的。

四、结论与展望

测试数据(2011年12月2日到2012年4月16日的79个数据)的标准差为751.6,表示测试数据的实际波动大小。用随机价差法得到的预测数据的波动性应该和实际波动性越接近越好,预测数据和实际数据的相关程度越高越好。表3给出了数值模拟的预测结果与实际值的相关系数和标准差的比较。结果显示,三种方法中,价差服从O-U过程的随机价差模型的预测值与实际值的相关系数最大,标准差最接近。基于Elliott框架的Shumway和Stoffer方法的预测值的标准差偏大,相关性最小;而Elliott和Krishnamurthy滤波方法的标准差偏小。产生这种结果的原因可以解释如下:Shumway和Stoffer方法中,“真实价差”的期望波动率是296.6,但是“真实价差”与观测数据的波动率较大,为600.8,这是使得预测数据的波动率比测试数据波动率大的原因;而Elliott和Krishnamurthy滤波方法中,“真实价差”的期望波动率为210.9,“真实价差”与观测数据的波动率较小,为190.9,使得预测数据的波动率比测试数据波动率小;而直接假定价差服从O-U过程的随机价差模型的期望波动率是679.85,与测试数据的波动率较为接近,所以本文认为预测最好的是价差服从O-U过程的随机价差模型。

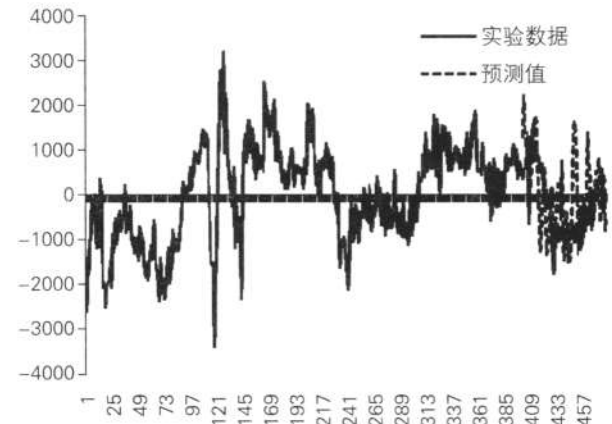


图3 Shumway和Stoffer方法的预测图

在利用协整理论设计套利策略时,通常使用价差spread的均值mean和标准差 σ 的组合作为开仓、平仓条件。例如若 $spread>mean+2\sigma$ 时,开仓卖出价差,等待 $spread<mean+0.2\sigma$ 时买入平仓。这样的策略是被动的,我们建仓后,就只能等待价差回复到均值水平再平仓套利,期间不能把握套利最大的时机以及无法预判价差回复到均值的时间。而本文讲述的随机价差法在协整套利策略的基础上,预测出未

来价差走势,在一定程度上克服了上述缺点。以图2描绘出的预测曲线为例,根据图中的预测结果,我们可以主动地选择最佳的建仓和平仓时机,从而使套利最大化。

我国股指期货和上证180ETF之间存在长期均值回复特性,为了介绍随机价差方法,本文选择O-U过程研究配对交易可行性,O-U过程是一类简单的均值回复过程,后续工作可以基于研究CIR过程等更复杂的均值回复过程的配对交易。



图4 Elliott和Krishnamurthy方法的预测图

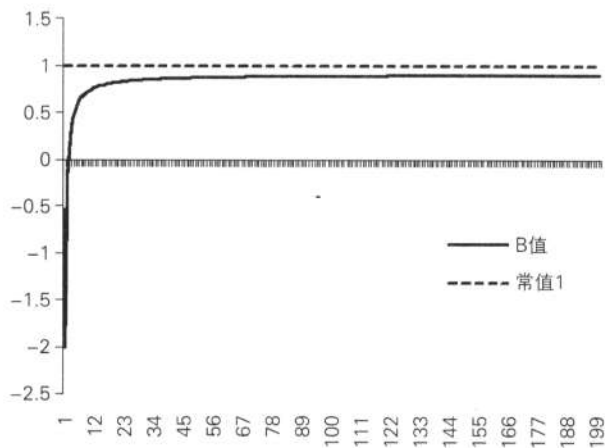


图5 Elliott和Krishnamurthy方法中B的收敛性

表3 原始数据和预测值的相关系数,预测值的标准差

	价差服从O-U过程	Shumway和Stoffer方法	Elliott和Krishnamurthy滤波方法
相关系数	0.536	0.464	0.522
标准差	791.72	919.15	574.53

参考文献:

[1]张敏,徐坚.ETF在股指期货期限套利的现货组合中的应用[J].技术经济与管理研究,2007,(3):31-32.

[2]薛战忠.股指期货与ETF的套利分析[J].北方经济,2011,(12):87-88.

[3]Gatev EG, Goetzmann WN, Rouwenhorst KG. Pairs Trading: Performance of a Relative Average Arbitrage Rule[J]. NBER Working Paper 7032.National Bureau of Economic Research Inc., 1999. Available online at: <http://www.nber.org/papers/w7032>.

[4]Feazzoetal et.al. Adapt the Basic Vasicek Interest-Rate Model to Generate a Mean Reversion Signal for Stock Pairs[J/OL]. Working Paper,2002. Available online at: <http://www. Insightful.com>.

[5]Vidyamurthy G. Pairs Trading: Quantitative Methods and Analysis[J].Hoboken NewJersey John Wiley&Sons.

[6]Elliott R J, Hoekd V D, Malcolm W P. Pairs Trading[J]. Quantitative Finance, 2005,3(5):271-276.

[7]Bertram W K. Analytic Solutions for Optimal Statistical Arbitrage Trading[J]. Physica A 389, 2010: 2234- 2243.

[8]Cummins M. Empirical Analysis of Optimal Statistical Arbitrage Trading on the Irish Stock Exchange[J/OL].2010, Available online at: <http://www. mark-cummins.com>.

[9]Finch S. Ornstein-Uhlenbeck Process[J/OL]. Unpublished Note. Available online at: <http://pauillac. inria.fr /algo /bsolve/ constant/constant.html>.

[10]Xiaoxia Feng, Dejun Xie. Bayesian Estimation of CIR Model[J]. Journal of Data Science,2012,10:271: 280.

[11]Dejun Xie. Parametric Estimation for Treasury Bills[J]. International Research Journal of Finance and Economics,2008,17:27-32.

[12]Shumway, R H and Stoffer, D S. An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm[J]. Journal of Time Series, 1982, 3:253-264.

[13]Elliott, R J and Krishnamurthy, V. Exact Finite-dimensional filters for Maximum Likelihood Parameter Estimation of Continuous-time Linear Gaussian Systems[J]. SIAM Journal of Control and Optimization, 1997, 35:1908-1923.

[14]Engle RF, Granger CWJ. Co-Integration and Error Correction: Representation, Estimation, and Testing[J]. Econometrica,1987,55(2):251-276.

(责任编辑:贾伟)