

## 基于 LSTM 神经网络的黑色金属期货套利策略模型

龙奥明<sup>1</sup>, 毕秀春<sup>2</sup>, 张曙光<sup>2</sup>

(1.中国科学技术大学数学科学学院, 安徽合肥 230026; 2.中国科学技术大学管理学院统计与金融系, 安徽合肥 230026)

**摘要:** 利用协整检验方法和 LSTM 神经网络算法, 建立黑色金属期货市场的套利策略模型. 利用基于 LSTM 神经网络套利策略模型对大连商品交易所上市的焦炭期货、铁矿石期货和上海期货交易所上市的螺纹钢期货进行实证研究. 对比研究基于 LSTM 神经网络、BP 神经网络和卷积神经网络的 3 种套利策略模型, 实证结果表明基于 LSTM 神经网络的黑色金属期货套利策略模型可行有效, 并且比 BP 神经网络套利策略模型和卷积神经网络套利策略模型表现更好.

**关键词:** 黑色金属期货; 跨品种套利; 协整; LSTM 神经网络

**中图分类号:** F830.9 **文献标识码:** A **doi:** 10.3969/j.issn.0253-2778.2018.02.006

**引用格式:** 龙奥明, 毕秀春, 张曙光. 基于 LSTM 神经网络的黑色金属期货套利策略模型[J]. 中国科学技术大学学报, 2018, 48(2): 125-132.

LONG Aoming, BI Xiuchun, ZHANG Shuguang. An arbitrage strategy model for ferrous metal futures based on LSTM neural network[J]. Journal of University of Science and Technology of China, 2018, 48(2): 125-132.

### An arbitrage strategy model for ferrous metal futures based on LSTM neural network

LONG Aoming<sup>1</sup>, BI Xiuchun<sup>2</sup>, ZHANG Shuguang<sup>2</sup>

(1. School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, China;

2. Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei 230026, China)

**Abstract:** Using the cointegration test method and LSTM neural network algorithm, the arbitrage strategy model for ferrous metal futures market was established. The empirical study is conducted on the coke futures, iron ore futures on the Dalian Commodity Exchange and the rebar futures on the Shanghai Futures Exchange using the arbitrage strategy model based on LSTM neural network. The arbitrage strategy models based on LSTM neural network, BP neural network and convolutional neural network were compared, and the empirical results show that the arbitrage strategy model for ferrous metal futures based on LSTM neural network is feasible and effective, and performs better than the arbitrage strategy models based on BP neural network and convolutional neural network.

**Key words:** ferrous metal futures; cross-commodity arbitrage; cointegration; LSTM neural network

收稿日期: 2017-05-23; 修回日期: 2017-11-09

基金项目: 国家自然科学基金(14401556, 14471304), 中央高校基本科研业务费专项资金(WK2040000012)资助.

作者简介: 龙奥明, 男, 1992 年生, 硕士. 研究方向: 金融工程. E-mail: aomingl@mail.ustc.edu.cn

通讯作者: 毕秀春, 博士/副研究员. E-mail: xcbi@ustc.edu.cn

## 0 引言

神经网络是人工智能领域的研究热点之一,在模式识别、智能机器人、自动控制、预测估计等领域广泛应用.神经网络的研究起源于 1943 年 McCulloch 和 Pitts 共同提出的单层感知器(MP)模型.经过不断深入研究,学者们提出了具有不同结构的神经网络模型.Hochreiter 和 Schmidhuber<sup>[1]</sup>在 1997 年提出了长短期记忆(long short term memory, LSTM)神经网络模型.Gers 等<sup>[2]</sup>在 2000 年通过引入遗忘门(forget gate)改进了 LSTM 神经网络模型.LSTM 神经网络在处理时间序列问题上取得了重要成果,现在广泛应用在语言系统<sup>[3]</sup>、图像识别与描述<sup>[4]</sup>、视频自动添加描述字幕<sup>[5]</sup>、机器翻译等领域.比如在 2016 年,谷歌将 LSTM 神经网络用于智能手机的智能助手 Allo 和谷歌翻译,苹果将 LSTM 神经网络用于 iPhone 的预测文本功能和智能助手 Siri.LSTM 神经网络对时间序列的强拟合性也可以运用到金融市场,如 Maknickienė 和 Maknickas<sup>[6]</sup>利用 LSTM 神经网络对欧元对美元汇率进行预测并建立交易策略,证明了基于 LSTM 神经网络的预测和交易策略是可行有效的;Persio 和 Honchar<sup>[7]</sup>利用 LSTM 神经网络对标普 500 指数的价格波动进行预测,结果证明了利用 LSTM 神经网络预测价格波动是可行有效的.

期货套利策略根据研究对象的不同可以分为期现套利、跨期套利、跨品种套利和跨市场套利.根据套利方法的不同,可分为根据统计学中如协整<sup>[8-10]</sup>、时间序列分析、随机控制<sup>[11]</sup>等分析方法的均值回归策略和基于神经网络算法的神经网络策略.目前国内外对神经网络套利策略的研究比较深入,如 Dunis 等<sup>[12]</sup>对比研究基于协整的均值回归策略和多种神经网络交易策略,对大豆、豆油之间的价差进行套利研究,结果表明神经网络策略优于均值回归策略;Dunis 等<sup>[13]</sup>利用多层感知器神经网络(MLP)、高阶神经网络(HONN)和遗传规划(GPA)算法对玉米、乙醇的压榨价差的套利策略进行了对比研究.而国内对神经网络套利策略的研究比较少,主要使用 BP 神经网络模型<sup>[14-16]</sup>.另外,文献<sup>[17-18]</sup>使用更新的卷积神经网络预测股市价格.考虑到 BP 神经网络具有收敛速度慢,预测精度不高等缺点和卷积神经网络作为前向神经网络在时间序列处理上的缺陷,因此我们可以使用更加高效的循环神经网络模型如

LSTM 神经网络用于研究期货市场的套利策略模型.由于铁矿石、焦炭和钢材是产业链的上下游关系,所以这 3 个品种之间很可能存在长期均衡关系(即协整关系).因此本文综合协整检验和 LSTM 神经网络,研究黑色金属期货市场上焦炭、铁矿石和螺纹钢期货 3 个品种之间的跨品种套利策略模型.

## 1 LSTM 神经网络套利策略模型的建立过程

### 1.1 LSTM 神经网络

循环神经网络(recurrent neural networks, RNNs)<sup>[19]</sup>是目前非常流行的神经网络模型.循环神经网络是专门用来处理时间序列问题的.对于前向神经网络(如 BP 神经网络)而言,从输入层到隐含层再到输出层,层与层之间是全连接的,每层之间的节点是无连接的,每一层的神经元信号只能够向下一层传播,样本的处理在时刻上是独立的,而对于循环神经网络而言,神经元在某时刻的输出可以直接影响该时刻的下一时刻的输入,因此循环神经网络能够处理时间序列方面的问题.

标准的 RNNs 学习从输入层序列  $(x_1, x_2, \dots, x_t)$  映射到隐藏状态序列  $(h_1, h_2, \dots, h_t)$  以及从隐藏状态序列  $(h_1, h_2, \dots, h_t)$  映射到输出层序列  $(z_1, z_2, \dots, z_t)$  的复杂的时空动态关系,映射过程基于以下公式:

$$h_t = f(W_{xh}x_t + W_{hh}h_{t-1} + b_h),$$

$$z_t = g(W_{hz}h_t + b_z).$$

式中,  $f, g$  为激励函数,通常设为 sigmoid 函数  $\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$  或 tanh 函数  $\text{tanh}(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ ;  $x_t, h_t \in \mathbb{R}^N$ ,  $z_t$  分别表示在  $t$  时刻的输入值、 $N$  个单元的隐藏状态值、输出值;  $W_{ij}$  为连接神经元层  $i, j$  的权重;  $b$  为偏倚.  $W_{ij}, b$  可以通过历史数据训练神经网络得到.

LSTM 神经网络是一种特殊的 RNNs,它可以解决普通 RNNs 的梯度消失和爆炸问题(即标准的 RNNs 结构难以传递相隔较远的信息).它在 RNNs 的基础上,加入了若干个控制数级的 gates,分别是输入门(input gates),遗忘门(forget gates),输出门(output gates),gate 的输出的值域为  $[0, 1]$ ,其中输入门控制有多少输入的信息可以流入记忆细胞(memory cell)中,遗忘门控制有多少上一时刻的记

忆细胞中的信息可以累积到当前时刻的记忆细胞中,输出门控制有多少当前记忆细胞的信息可以流入到当前隐藏状态中,其结构如图 1 所示。

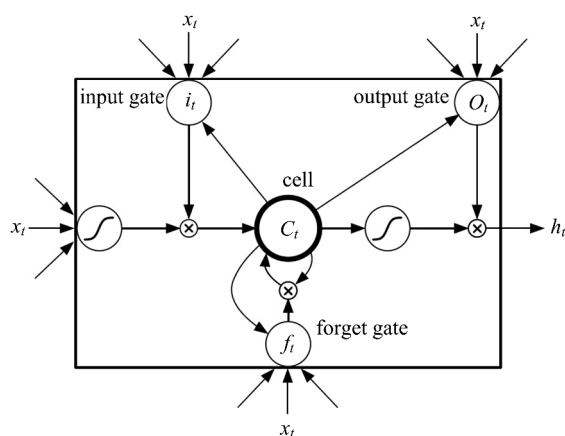


图 1 LSTM 神经网络结构图

Fig.1 Structure of LSTM neural network

LSTM 神经网络的数学公式如下:

input gates:

$$i_t = \text{sigmoid}(W_{xi}x_t + W_{hi}h_{t-1} + b_i).$$

forget gates:

$$f_t = \text{sigmoid}(W_{xf}x_t + W_{hf}h_{t-1} + b_f).$$

output gates:

$$o_t = \text{sigmoid}(W_{xo}x_t + W_{ho}h_{t-1} + b_o).$$

cells:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c).$$

cell outputs:

$$h_t = o_t \odot \tanh(c_t).$$

相比于 RNNs, LSTM 神经网络多出的参数  $i_t$ ,  $f_t$ ,  $o_t$ ,  $c_t$  分别表示在  $t$  时刻输入门的输出值、遗忘门的输出值、输出门的输出值、记忆细胞状态;  $\odot$  表示元素积 (element-wise product)。

## 1.2 协整

1987 年 Engle 和 Granger 提出的协整理论及其方法<sup>[20]</sup>为非平稳序列的建模提供了一种途径。将  $d$  阶单整序列记为  $I(d)$ , 如果  $k$  维向量  $Y_t$  的每个分量都是  $d$  阶单整  $I(d)$ , 且至少存在它们的一个线性组合服从  $d-c$  阶单整, 其中  $d \geq c > 0$ , 则称向量  $Y_t$  的元素之间满足  $d, c$  阶协整, 记为  $Y_t \sim CI(d, c)$ 。即  $Y_t$  的分量间存在协整关系当且仅当如下关系成立:

$$\beta'_i Y_t = z_{i,t} \sim I(d-c), i=1, \dots, r.$$

式中,  $\beta_i$  称为协整向量。

验证三个价格序列存在协整关系的步骤如下:

## 第一步: 平稳性检验

分别对三个序列进行平稳性检验, 如果序列存在单位根, 则说明序列不平稳, 需要对其进行一阶差分, 再对差分后的序列进行平稳性检验, 如果满足平稳性检验, 转入下一步。

## 第二步: 协整检验

如果三个序列都服从相同阶数的单整, 则检验序列之间是否存在协整关系, 即对该三个序列建立线性回归方程, 若残差不平稳, 则三个序列之间不存在协整关系; 若残差平稳, 则三个序列之间存在协整关系。

## 1.3 套利策略模型

用  $\text{rebar}_t, \text{coke}_t, (\text{iron\_ore})_t$  分别表示  $t$  时刻的螺纹钢、焦炭和铁矿石期货的价格, 如果它们三者满足 1, 1 阶协整关系, 则它们可以由如下线性回归方程所表示:

$$\text{rebar}_t = a \times \text{coke}_t + b \times (\text{iron\_ore})_t + c + \varepsilon_t.$$

因此可以根据螺纹钢、焦炭和铁矿石期货价格满足的线性回归方程的系数  $a, b$  对不同期货品种进行配比建仓, 得到产业链的“冶炼价值”spread 的表达式:

$$\text{spread}_t = \text{rebar}_t - a \times \text{coke}_t - b \times (\text{iron\_ore})_t.$$

交易头寸设置: 将  $\text{spread}_t$  作为套利策略模型的交易头寸。买入  $\text{spread}_t$ , 相当于买入  $\text{rebar}_t$ , 同时卖出  $a \times \text{coke}_t$  和  $b \times (\text{iron\_ore})_t$ ; 卖出  $\text{spread}_t$ , 相当于卖出  $\text{rebar}_t$ , 同时买入  $a \times \text{coke}_t$  和  $b \times (\text{iron\_ore})_t$ 。

### 1.3.1 LSTM 神经网络预测及评价

本文用到的神经网络都是使用 Python 里面的 Keras 深度学习框架。我们使用改进的 LSTM 神经网络进行预测, 它包括 2 个 LSTM 层和 1 个全连接层, 其结构如图 2 所示。

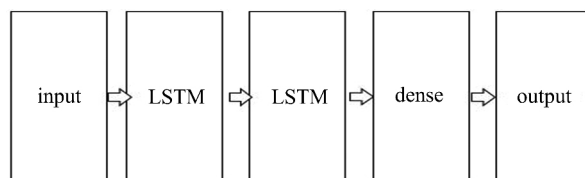


图 2 改进的 LSTM 神经网络结构

Fig.2 Improved LSTM neural network structure

将历史数据分成训练集和测试集两部分, 利用训练集的数据对改进的 LSTM 神经网络进行训练, 再对测试集的数据进行预测。定义  $\text{spread}_t$  的增量为

$$\Delta \text{spread}_t = \text{spread}_t - \text{spread}_{t-1}.$$

利用改进的 LSTM 神经网络对  $\text{spread}_t$  进行非线性拟合,即用  $t$  时刻之前的连续  $T$  个  $\text{spread}$  数据(不包括  $t$  时刻)作为输入值去预测输出值  $\text{spread}_t$ ,设得到的预测值为  $\text{predict\_spread}_t$ ,因此我们得到预测的增量值为

$$\Delta s\_pred_t = \text{predict\_spread}_t - \text{spread}_{t-1}.$$

对预测结果建立如下两个指标对预测结果进行评价:

平均绝对误差(MAE):

$$\text{MAE} = \frac{\sum_{i=1}^N |\text{predict\_spread}_i - \text{spread}_i|}{N}.$$

符号正确率(SA):

$$\text{SA} = \frac{\sum_{i=1}^N z_i}{N}.$$

式中,  $\Delta s\_pred_t \times \Delta s_t > 0$  时  $z_t = 1$ , 否则  $z_t = 0$ .

### 1.3.2 交易策略

根据神经网络的预测值我们可以建立交易策略,当预测的  $\text{spread}$  的增量的绝对值大于一个阈值(分固定和浮动两种情况)时,进行开仓交易,当预测到  $\text{spread}$  将反向变化时进行平仓.考虑到螺纹钢、铁矿石和焦炭的价格满足协整关系,则  $\text{spread}$  具有均值回复的特征,可以在开仓条件中加上  $\text{spread}$  处于在均值  $\text{mean}(\text{spread})$  的两倍标准差  $2\text{stv}(\text{spread})$  之外的限制条件.故得到以下 4 种交易策略:

交易策略 1(固定阈值):

①如果  $\Delta s\_pred_t > X$ , 则买入  $\text{spread}_{t-1}$  开仓,当  $\Delta s\_pred_t < 0$  时平仓;

②如果  $\Delta s\_pred_t < -X$ , 则卖出  $\text{spread}_{t-1}$  开仓,当  $\Delta s\_pred_t > 0$  时平仓;

交易策略 2(浮动阈值):

①如果  $\Delta s\_pred_t > X(t)$ , 则买入  $\text{spread}_{t-1}$  开仓,当  $\Delta s\_pred_t < 0$  时平仓;

②如果  $\Delta s\_pred_t < -X(t)$ , 则卖出  $\text{spread}_{t-1}$  开仓,当  $\Delta s\_pred_t > 0$  时平仓;

交易策略 3(固定阈值+均值回复):

①如果  $\Delta s\_pred_t > X$  且  $\text{spread}_{t-1} < \text{mean}(\text{spread}) - 2\text{stv}(\text{spread})$ , 则买入  $\text{spread}_{t-1}$  开仓,当  $\Delta s\_pred_t < 0$  时平仓;

②如果  $\Delta s\_pred_t < -X$  且  $\text{spread}_{t-1} > \text{mean}(\text{spread}) + 2\text{stv}(\text{spread})$ , 则卖出  $\text{spread}_{t-1}$  开仓,当  $\Delta s\_pred_t > 0$  时平仓;

交易策略 4(浮动阈值+均值回复):

①如果  $\Delta s\_pred_t > X(t)$  且  $\text{spread}_{t-1} < \text{mean}(\text{spread}) - 2\text{stv}(\text{spread})$ , 则买入  $\text{spread}_{t-1}$  开仓,当  $\Delta s\_pred_t < 0$  时平仓;

②如果  $\Delta s\_pred_t < -X(t)$  且  $\text{spread}_{t-1} > \text{mean}(\text{spread}) + 2\text{stv}(\text{spread})$ , 则卖出  $\text{spread}_{t-1}$  开仓,当  $\Delta s\_pred_t > 0$  时平仓.

为了对 LSTM 神经网络套利策略模型的实证结果进行对比分析,本文使用国内应用最多的 BP 神经网络和更新的卷积神经网络替换改进的 LSTM 神经网络,就可以得到 BP 神经网络的套利策略模型和卷积神经网络套利策略模型.它们的结构如图 3、图 4 所示.由于理论上具有两个全连接层的 BP 神经网络可以拟合所有的非线性情况,因此我们使用具有 2 个全连接层的 BP 神经网络.

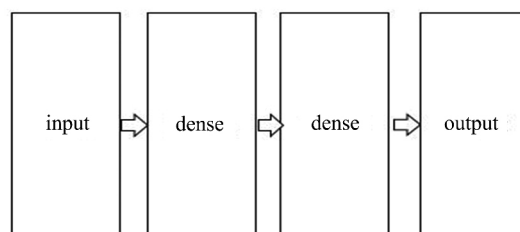


图 3 BP 神经网络结构

Fig.3 BP neural network structure

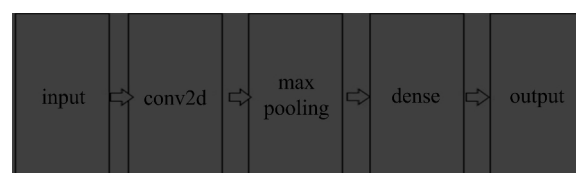


图 4 卷积神经网络结构

Fig.4 Convolutional neural network structure

## 2 实证分析

### 2.1 数据选取

选取大连商品交易所上市的焦炭期货、铁矿石期货和上海期货交易所上市的螺纹钢期货为研究对象.由于市场上存在交割月份不同的期货合约,为了覆盖掉流动性风险,这里选取持仓量和成交量最大的合约(即主力合约)的收盘价构建该品种期货的连续收盘价数据.主力连续合约历史数据可以从期货交易软件开拓者 TB 下载得到.由于铁矿石期货 2013 年 10 月 18 日才在大连商品交易所上市,而且各个品种夜盘交易的开放时间不同,交易时间也有所差别.因此选取 2013-10-18~2017-03-31 的主

力连续合约日盘的历史行情的每 5 min 收盘价的高频数据作为实证研究数据。

焦炭、铁矿石和螺纹钢期货的收盘价时间序列如图 5 所示,可以看出它们走势相近,它们之间可能存在某种稳定的关系。

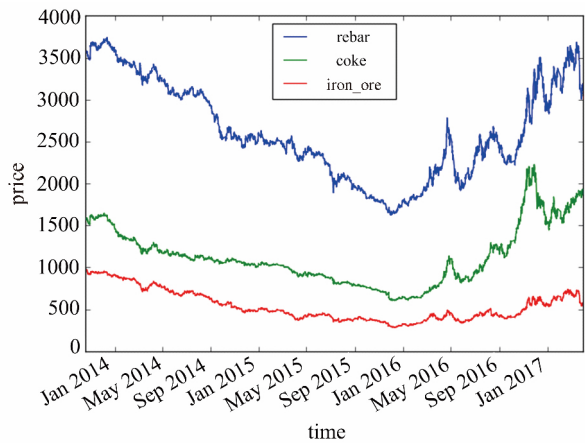


图 5 焦炭、铁矿石和螺纹钢期货的收盘价时间序列  
Fig.5 The time series of the closing price of coke, iron ore and rebar futures

2.2 协整检验

在进行协整检验之前,对焦炭、铁矿石和螺纹钢期货的价格序列进行 ADF 检验,结果如表 1 所示.其中 coke,iron\_ore,rebar 分别表示焦炭、铁矿石和

螺纹钢期货的价格序列; $\Delta\text{coke}$ , $\Delta\text{iron\_ore}$ , $\Delta\text{rebar}$  分别表示焦炭、铁矿石和螺纹钢期货的价格序列的一阶差分.根据检验结果,在 1%,5%,10% 的显著性水平上,coke,iron\_ore,rebar 的 ADF 检验值大于临界值,故它们都是非平稳序列;一阶差分序列  $\Delta\text{coke}$ , $\Delta\text{iron\_ore}$ , $\Delta\text{rebar}$  的 ADF 检验值小于临界值,故一阶差分序列都是平稳序列.则 coke,iron\_ore,rebar 都是一阶单整  $I(1)$  过程,可以进一步进行协整检验.

利用最小二乘法建立 coke,iron\_ore,rebar 的线性回归方程,以 rebar 为因变量,得到回归方程为

$$\begin{aligned} \text{rebar}_t = & 0.5173 \times \text{coke}_t + \\ & (0.0000) \\ & 2.3716 \times (\text{iron\_ore})_t + 766.919 + \epsilon_t. \\ & (0.0000) \quad (0.0000) \end{aligned}$$

其中,  $R^2=0.9548$ , 括号内为对应的伴随概率,可以看出拟合程度较好.则可以得到如下的残差序列的表达式:

$$\begin{aligned} \epsilon_t = & \text{rebar}_t - 0.5173 \times \text{coke}_t - \\ & 2.3716 \times (\text{iron\_ore})_t - 766.919. \end{aligned}$$

对残差 resid 进行 ADF 单位根检验,结果如表 2 所示.且可以得到残差的均值为  $1.93\text{E}-13$ ,标准差为 123.1791.

表 1 焦炭、铁矿石和螺纹钢期货的价格序列及其一阶差分的 ADF 检验结果

Tab.1 The ADF test results of coke, iron\_ore, rebar futures and their first order difference

| 序列                       | ADF 检验值    | 1%临界值      | 5%临界值      | 10%临界值     | 伴随概率    |
|--------------------------|------------|------------|------------|------------|---------|
| coke                     | -0.385 333 | -3.430 352 | -2.861 425 | -2.566 749 | 0.909 4 |
| iron_ore                 | -2.550 535 | -3.430 352 | -2.861 425 | -2.566 749 | 0.103 6 |
| rebar                    | -1.496 711 | -3.430 352 | -2.861 425 | -2.566 749 | 0.535 5 |
| $\Delta\text{coke}$      | -199.640 1 | -3.430 352 | -2.861 425 | -2.566 749 | 0.000 1 |
| $\Delta\text{iron\_ore}$ | -143.974 3 | -3.430 352 | -2.861 425 | -2.566 749 | 0.000 1 |
| $\Delta\text{rebar}$     | -197.148 6 | -3.430 352 | -2.861 425 | -2.566 749 | 0.000 1 |

表 2 残差的 ADF 检验结果

Tab.2 ADF test results of residuals

| 序列    | ADF 检验值    | 1%临界值      | 5%临界值      | 10%临界值     | 伴随概率    |
|-------|------------|------------|------------|------------|---------|
| resid | -3.763 735 | -3.430 352 | -2.861 425 | -2.566 749 | 0.003 3 |

残差的 ADF 检验值 -3.763735 小于 1%置信水平下的临界值 -3.430352,因此可以认为在 1%置信水平下,残差序列是平稳的.故焦炭、铁矿石和螺纹钢期货的价格满足协整关系,三者之间存在长期均衡关系。

2.3 LSTM 神经网络策略与 BP 神经网络策略实证分析

根据焦炭、铁矿石和螺纹钢期货的价格的线性回归表达式,对不同期货品种配比建仓,可以得到关于钢铁行业产业链“冶炼价值”spread 的时间序列

和 spread 的增量的时间序列,其计算公式分别如下:

$$\begin{aligned} \text{spread}_t &= \text{rebar}_t - 0.5173 * \text{coke}_t - \\ &\quad 2.316 * (\text{iron\_ore})_t, \\ \text{Delta\_s}_t &= \text{spread}_t - \text{spread}_{t-1}. \end{aligned}$$

这里用  $t$  时刻之前的连续 10 个 spread 数据 (不包括  $t$  时刻) 去预测  $\text{spread}_t$ , 设得到的预测值为  $\text{predict\_spread}_t$ .

$$\text{Delta\_s\_pred}_t = \text{predict\_spread}_t - \text{spread}_{t-1}.$$

取 2013-10-18~2016-09-06 共 31941 个滚动分组的数据作为神经网络样本内的训练集, 2016-09-07 日~2017-03-31 共 7075 个数据作为神经网络样本外的测试集.

为了排除神经网络参数选取的不同对实证结果的影响, 两个 LSTM 层神经元个数分别从 50 到 100 每间隔 5 进行一次取值, 全连接层节点数从 10 到 20 每间隔 1 进行一次取值, 总共形成  $11^3$  个不同参数的神经网络. 训练使用 RMSProp 优化算法, 损失函数使用均方误差 MSE, 训练 LSTM 神经网络使损失函数最小. 选取损失函数收敛到最小的神经网络作为最优的 LSTM 神经网络, 记录下其参数为 (60, 65, 16).

再将非线性拟合所用的 LSTM 神经网络替换成 BP 神经网络和卷积神经网络. 利用类似的方法得到最优的 BP 神经网络的参数 (全连接层节点数, 全连接层节点数) 为 (16, 16), 最优的卷积神经网络的参数 (卷积核大小, 全连接层节点数) 为 (1×8, 16).

则得到在 2016-09-07~2017-03-31 的时间区间上 3 种神经网络的预测结果分别如图 6~8 所示, 预测评价指标如表 3 所示.

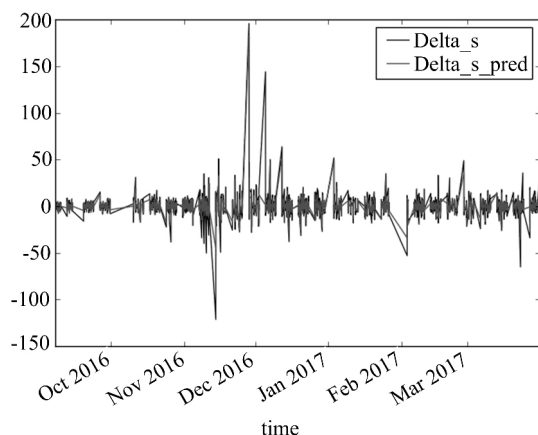


图 6 LSTM 神经网络预测结果

Fig.6 LSTM neural network prediction results

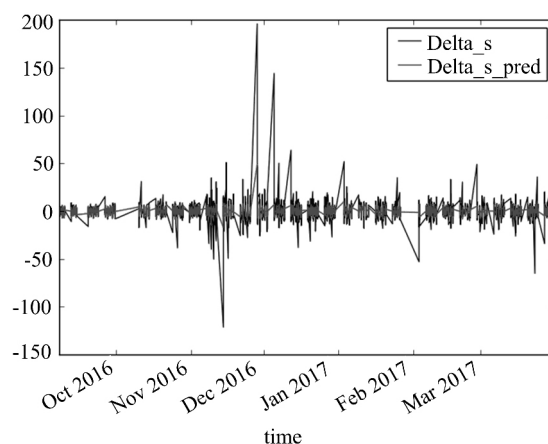


图 7 BP 神经网络预测结果

Fig.7 BP neural network prediction results

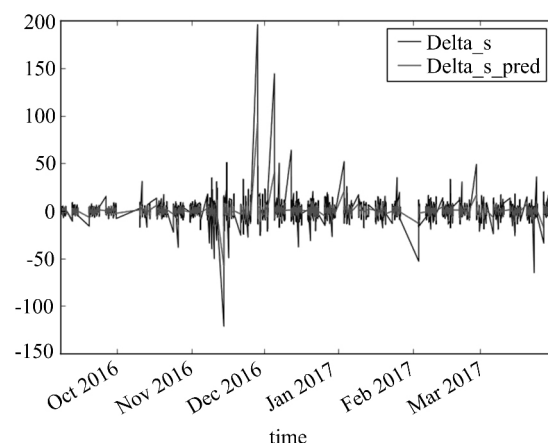


图 8 卷积神经网络预测结果

Fig.8 Convolutional neural network prediction results

表 3 3 种神经网络预测评价

Tab.3 Evaluation of the three neural network prediction results

| 评价指标       | LSTM 神经网络 | BP 神经网络 | 卷积神经网络 |
|------------|-----------|---------|--------|
| 平均绝对误差 MAE | 2.42      | 3.79    | 3.08   |
| 符号正确率 SA   | 75.9%     | 55.2%   | 63.7%  |

从图 6~8 可以看出: LSTM 神经网络的预测结果最好, 卷积神经网络次之, BP 神经网络的预测结果不太理想. 这是因为 LSTM 神经网络是循环神经网络, 可以将  $t$  时刻的输出反馈到下一个时刻的输入中去, 它能够较好地提取时间序列的信息, 而 BP 神经网络是前向神经网络, 它在每个时刻上的处理是独立的, 这与时间序列的情况不符, 从而预测结果不太好. 卷积神经网络虽然也是前向神经网络, 但是它可以提取时间序列的局部特征, 故表现得比 BP 神经网络更优秀. 因此, LSTM 神经网络在预测时间序列上更具有优势.

从表3可以看出:无论是从平均绝对误差还是符号正确率上,LSTM神经网络都比BP神经网络和卷积神经网络更优秀.BP神经网络预测的符号正确率只有55%左右,卷积神经网络可以达到60%左右,而LSTM神经网络预测的符号正确率可以达到75%以上,这说明只要选择合适的交易条件,LSTM神经网络的套利策略是可行的.

由于螺纹钢交易单位是10吨/手,焦炭和铁矿石交易单位是100吨/手,而它们的报价单位是元/吨.因此,买入spread相当于买入10手螺纹钢期货,同时卖出0.5173手焦炭期货和2.3716手铁矿石期货;卖出spread相当于卖出10手螺纹钢期货,同时买入0.5173手焦炭期货和2.3716手铁矿石期货.实际交易中交易手数必须是整数,则同比例放大三者的头寸即可.通过查找大连商品交易所和上海期货交易所提供的信息,焦炭、铁矿石和螺纹钢期货的交易手续费分别为成交金额的万分之一点二、万分之一点二和万分之一,但实际的交易手续费要比交易所提供的更高,在本文的研究中,设定交易手续费为成交金额的万分之二点五,计算历史数据得到交易一份头寸的手续费为1.31~3.42元/吨.

为了使每次交易盈利覆盖掉手续费,我们取固

定阈值 $X$ 为7,浮动阈值 $X(t)$ 为 $t$ 时刻之前的连续10个交易日Bar的手续费的均值的3倍.交易头寸spread的均值 $\text{mean}(\text{spread})=766.919$ ,标准差 $\text{stv}(\text{spread})=123.1791$ .

根据预测结果和给定的交易策略,则得到在2016-09-07~2017-03-31的时间区间上3种神经网络套利策略的交易绩效结果,如表4所示.从表4可以看出:在相同策略下,3种神经网络都有较高的胜率,LSTM最高,卷积神经网络次之,BP最低;且LSTM的总收益和平均单笔收益都明显多于BP和卷积神经网络.这说明LSTM神经网络的胜率、总收益和单笔收益都优于BP和卷积神经网络,当然卷积神经网络优于BP神经网络.相同神经网络情况下,从策略1到策略4,胜率递增.策略1、2的胜率明显小于策略3、4,交易次数明显多于策略3、4,平均单笔收益明显小于策略3、4.这是因为策略1、2具有更多的开仓次数,需要缴纳的手续费显著增多,因此单笔收益更低.所有模型中,基于LSTM神经网络的策略4模型可以取得最优的胜率和单笔收益率,但是基于LSTM神经网络的策略2模型可以取得最大总收益.

表4 3种神经网络套利策略交易绩效结果

Tab.4 Trading performance results of the three neural network arbitrage strategies

| 交易策略     | LSTM 神经网络 |         |        |        | BP 神经网络 |         |        |        | 卷积神经网络  |         |        |        |
|----------|-----------|---------|--------|--------|---------|---------|--------|--------|---------|---------|--------|--------|
|          | 策略1       | 策略2     | 策略3    | 策略4    | 策略1     | 策略2     | 策略3    | 策略4    | 策略1     | 策略2     | 策略3    | 策略4    |
| 开仓次数     | 738       | 852     | 55     | 57     | 614     | 723     | 42     | 44     | 702     | 787     | 45     | 48     |
| 盈利次数     | 539       | 647     | 43     | 46     | 387     | 470     | 29     | 30     | 491     | 567     | 34     | 37     |
| 胜率       | 0.73      | 0.76    | 0.78   | 0.80   | 0.63    | 0.65    | 0.68   | 0.70   | 0.70    | 0.72    | 0.76   | 0.77   |
| 总收益/元    | 305 746   | 393 794 | 47 079 | 50 565 | 166 344 | 227 015 | 31 617 | 31 717 | 281 418 | 355 677 | 32 437 | 35 250 |
| 平均单笔收益/元 | 414.29    | 462.2   | 855.98 | 887.11 | 270.92  | 313.99  | 752.79 | 720.84 | 400.88  | 451.94  | 720.82 | 734.38 |

### 3 结论

本文证明了黑色金属期货中的焦炭、铁矿石和螺纹钢期货价格之间存在长期均衡关系.根据三者价格满足的线性回归方程,构造三者之间的投资组合作为套利的头寸,利用LSTM神经网络对头寸进行非线性拟合,进而建立基于LSTM神经网络的套利策略模型.实证结果表明,LSTM神经网络套利策略模型可行有效.将LSTM神经网络替换成常见的BP神经网络和更新的卷积神经网络之后,可以发

现,LSTM神经网络套利策略无论是预测的精确度方面还是投资收益方面都明显优于BP神经网络套利策略和卷积神经网络套利策略.

本文的创新点在于把主要运用在语言系统、图像识别与描述、机器翻译等领域的LSTM神经网络运用到黑色金属期货市场并建立套利策略模型.本文基于协整理论建立跨品种的头寸,基于LSTM神经网络预测的头寸值设置套利交易策略模型,通过对头寸的增加量设定固定阈值和随时间变化的浮动阈值以及是否利用协整关系的均值回复特征设立4

种不同的交易策略,通过与 BP 神经网络和卷积神经网络实证对比,证明了 LSTM 神经网络在套利交易模型中具有的优势.后续可以从人工神经网络理论创新的角度继续深入研究.

#### 参考文献(References)

- [1] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [2] GERS F A, SCHMIDHUBER J, CUMMINS F. Learning to forget: Continual prediction with LSTM[J]. *Neural Computation*, 2000, 12(10): 2451-2471.
- [3] WEN T H, GASIC M, MRKSIC N, et al. Semantically conditioned LSTM-based natural language generation for spoken dialogue systems [C]// *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Stroudsburg, PA: Association for Computational Linguistics, 2015: 1711-1721.
- [4] DONAHUE J, HENDRICKS A L, GUADARRAMA S, et al. Long-term recurrent convolutional networks for visual recognition and description[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2015: 2625-2634.
- [5] VENUGOPALAN S, XU H, DONAHUE J, et al. Translating videos to natural language using deep recurrent neural networks [J]. *Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL*. Stroudsburg, PA: Association for Computational Linguistics, 2015: 1494-1504.
- [6] MAKNIKENÉ N, MAKNICKAS A. Application of neural network for forecasting of exchange rates and forex trading [C]// *The 7th International Scientific Conference "Business and Management 2012"*. Vilnius, Lithuania: Vilnius Gediminas Technical University, 2012: 10-11.
- [7] DI PERSIO L, HONCHAR O. Artificial neural networks architectures for stock price prediction: Comparisons and applications[J]. *International Journal of Circuits, Systems and Signal Processing*, 2016, 10: 403-413.
- [8] SIMON D P. The soybean crush spread: Empirical evidence and trading strategies[J]. *Journal of Futures Markets*, 1999, 19(3): 271-289.
- [9] 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.  
QIU Zhongqun, CHENG Xijun. Calendar spread arbitrage strategy model for index futures based on co-integration rule[J]. *Systems Engineering*, 2008, 26(12): 26-29.
- [10] 葛翔宇, 吴洋, 周艳丽. 门限协整套利: 理论与实证研究[J]. *统计研究*, 2012, 29(3): 79-87.  
GE Xiangyu, WU Yang, ZHOU Yanli. Threshold cointegration arbitrage: Theory and application [J]. *Statistical Research*, 2012, 29(3): 79-87.
- [11] KANAMURA T, RACHEV S T, FABOZZI F J. A profit model for spread trading with an application to energy futures[J]. *The Journal of Trading*, 2010, 5(1): 48-62.
- [12] DUNIS C L, LAWS J, EVANS B. Modelling and trading the soybean-oil crush spread with recurrent and higher order networks: A comparative analysis [J]. *Neural Network World*, 2006, 16(3): 193.
- [13] DUNIS C L, LAWS J, MIDDLETON P W, et al. Trading and hedging the corn/ethanol crush spread using time-varying leverage and nonlinear models[J]. *The European Journal of Finance*, 2015, 21(4): 352-375.
- [14] 曾谦, 马丹頔, 刘宗鑫. 基于 BP 神经网络改进的黄金价格预测[J]. *计算机仿真*, 2010(9): 200-203.  
ZENG Lian, MA Dandi, LIU Zongxin. Gold price forecast based on improved BP neural network [J]. *Computer Simulation*, 2010, 27(9): 200-203.
- [15] 张金仙, 闫二乐, 杨拴强. 基于自适应 BP 神经网络的上证指数预测模型的研究[J]. *长春大学学报*, 2016, 26(6): 26-30.  
ZHANG Jinxian, YAN Erle, YANG Shuanqiang. Research on prediction model of shanghai stock exchange index based on self-adaptive BP neural network [J]. *Journal of Changchun University*, 2016, 26(6): 26-30.
- [16] 林杰, 龚正. 基于人工神经网络的沪锌期货价格预测研究[J]. *财经理论与实践*, 2017, 38(2): 54-57.  
LING Jie, GONG Zheng. A research on forecasting of Shanghai zinc futures price based on artificial neural network [J]. *The Theory and Practice of Finance and Economics*, 2017, 38(2): 54-57.
- [17] 张贵勇. 改进的卷积神经网络在金融预测中的应用研究[D]. 郑州: 郑州大学, 2016.  
ZHANG Guiyong. Research on the application of improved convolutional neural network in financial forecasting [D]. Zhengzhou: Zhengzhou University, 2016.
- [18] TSANTEKIDIS A, PASSALIS N, TEFAS A, et al. Forecasting stock prices from the limit order book using convolutional neural networks [C]// *2017 IEEE 19th Conference on Business Informatics*. IEEE, 2017: 7-12.
- [19] GRAVES A. Supervised Sequence Labelling with Recurrent Neural Networks [M]. Berlin: Springer, 2012: 15-35.
- [20] ENGLE R F, GRANGER C W J. Co-integration and error correction: Representation, estimation, and testing[J]. *Econometrica: Journal of the Econometric Society*, 1987, 55(2): 251-276.