

文章编号: 1671-4229(2025)02-0065-09

基于相空间重构和门控循环单元的股价预测研究

刘广¹, 伍子睿², 黄思杰^{3*}

(1. 广州大学 经济与统计学院, 广东 广州 510006; 2. 香港浸会大学 工商管理学院, 香港 999077;
3. 广州华立学院 商学院, 广东 广州 511325)

摘要: 准确预测股价波动是金融投资关注的焦点问题。股价波动受多种因素影响, 具有非线性特征, 传统的线性预测方法往往难以奏效。文章选择 A 股 5 个代表性股指与 5 只大市值股票为样本, 使用其 2020—2023 年的日收盘价数据, 首先, 借助相空间重构技术 (Phase Space Reconstruction, PSR) 将股价时间序列映射到高维空间中, 揭示其混沌特征; 然后, 基于门控循环单元 (Gate Recurrent Unit, GRU) 深度学习方法开发出 PSR-GRU 预测模型, 生成股价预测结果; 最后, 将预测结果与经典预测模型所得结果进行对比。结果发现, 股价波动具有混沌特性, PSR-GRU 模型在股价预测上表现出更优异的性能。

关键词: 相空间重构; 门控循环单元; 混沌理论; 深度学习

中图分类号: F061.1 文献标志码: A

Analysis of the stock price prediction based on phase space reconstruction and gated recurrent unit

LIU Guang¹, WU Zi-rui², HUANG Si-jie^{3*}

(1. School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China;
2. School of Business, Hong Kong Baptist University, Hong Kong 999077, China;
3. School of Commerce, Guangzhou Huali College, Guangzhou 511325, China)

Abstract: It is critical to predict stock price fluctuations accurately in financial investment. Stock price fluctuations are influenced by multiple factors and have nonlinear characteristics, making traditional linear prediction methods often ineffective. We first selected 5 representative stock indices and 5 large-cap stocks of A-shares as samples, used their daily closing price data from 2020 to 2023, to map their price time series to a high-dimensional space by using the phase space reconstruction (PSR) technique, and reveal their chaotic characteristics. Then, based on the deep learning method Gated Recurrent Unit (GRU), we developed a PSR-GRU prediction method to generate stock price prediction results. Finally, we compared the predicted results with those obtained from classical prediction models. We found that stock price fluctuations have chaotic characteristics, and the PSR-GRU exhibits superior performance in stock price prediction.

Key words: phase space reconstruction; gate recurrent unit; chaos theory; deep learning

收稿日期: 2024-09-25; 修回日期: 2025-02-18

作者简介: 刘广 (1980—), 男, 副教授, 博士. E-mail: mrluoguang@126.com

* 通信作者. E-mail: S. Huang. Roy@outlook.com

引文格式: 刘广, 伍子睿, 黄思杰. 基于相空间重构和门控循环单元的股价预测研究[J]. 广州大学学报(自然科学版), 2025, 24(2): 65-73.

股价波动是金融投资中关注的焦点问题,股价预测因此成为金融领域研究的核心主题。传统观点认为,股价运动服从随机游走,波动无序且不可测。此时证券收益率服从独立正态分布,市场强势有效,没有“免费午餐”。但实际上,股价波动并非随机的,具有短期自相关性,收益率分布呈现尖峰肥尾特征^[1]。传统的时间序列模型,如差分自回归移动平均模型,虽然在处理线性序列数据时应用广泛,但在揭示非线性特征方面力有不逮。

随着非线性科学发展,越来越多的研究开始关注股价波动的非线性特性,并尝试揭示其运动规律。有研究表明,股价序列具有混沌特征,受初值敏感影响较大,导致长期预测难度较高^[2]。相空间重构技术(PSR)为探索和预测股价波动的混沌特性提供了新的视角,在金融领域的应用正逐步扩大,逐渐成为预测股价走势的重要手段。该技术通过从一维时间序列中提取特征并将其映射到高维空间,揭示了数据的内部动态结构,从而增强了对时序数据的理解和分析能力。

深度学习方法较多,主要包括支持向量机(Support Vector Machine, SVM)、循环神经网络(Recurrent Neural Network, RNN)等。SVM具有强大的非线性建模能力,但需要假设样本之间独立,导致适应性不足。RNN无需假设样本独立,并具备记忆功能,适合处理具有时间关联的序列数据。但RNN在训练中可能遭遇梯度消失或爆炸问题,难以捕捉时间序列的长期依赖性。其后,基于RNN发展出长短期记忆网络(Long Short-Term Memory, LSTM)和门控循环单元(GRU)。LSTM包含3个门控和1个记忆单元,能有效捕捉长期依赖关系。GRU结构更简单,只包含更新门和重置门,参数更少,训练更快,适用于需快速迭代的场景,逐渐成为处理复杂序列数据的首选方法。

本文构建PSR-GRU模型对股价数据进行分析和预测,主要贡献在于:①揭示了股价序列数据的复杂结构;②提供了一种适当的预测方法,该方法不仅适用于股价数据,还适用于其他具有混沌性质的时序数据,为理解与预测复杂现象提供了新的视角和工具。

1 文献综述

PSR由Packard等^[3]首次提出,假设复杂系统

中任一变量的变化都包含其他变量的信息,它已被广泛应用于多种不同领域的混沌时间序列研究中。其后,Lang等^[4]通过多种算法分析混沌时间序列的动态特性,给出最优参数确定方法,极大提高了模型预测性能。

近年来,金融系统的混沌特性受到学界的极大关注,诸多研究据此实施股价预测。Dousseh等^[5]探讨了整数和分数阶金融系统的混沌控制问题,发现混沌行为可以被控制到系统的任何不稳定平衡点。Huang等^[6]通过流形学习技术探索了中国金融市场的层次结构,发现市场的混沌程度相对较低,但脆弱性高,在危机中更易于崩溃。Shah等^[7]在分数阶框架中构建一种金融系统模型,分析了其混沌行为。这些研究体现了在金融领域应用先进的数学和统计方法来探索和解释市场动态行为的趋势,特别有助于理解市场在各种经济力量作用下如何表现出非线性,从而深化了对金融市场混沌特征的理解。Zhang等^[8]将PSR与回声状态网络相结合,证明其在股票价格预测中是有效和高效的。Khuwaja等^[9]使用PSR和极值学习机预测股票价格走势,极大提高了预测性能。Li等^[10]提出一种基于路径签名的PSR方法用于股票趋势预测,部分解决了仅依据PSR方法时面临的矩阵维数高、预测过拟合和计算效率低的问题。我国金融市场的混沌特征也被研究所证实^[11-14]。陈佐等^[15]将一种基于PSR的期望最大化聚类模糊神经网络预测模型应用于股价预测,发现预测模型降低了预测误差,提高了系统的性能。郭兆纲等^[16]基于PSR提出边缘点分析方法,发现能够探测影响金融市场任一时间区域的重大历史事件。

深度学习最早由Hinton等^[17]提出,近年广泛应用于金融预测领域^[18]。Jin等^[19]开发了一种基于经验模态分解和反向传播神经网络的深度学习模型,展示出卓越的预测性能。马甜等^[20]将深度学习模型应用于中国股市,发现收益预测精度和因子投资绩效均有很大提升。

随着非线性时间序列研究的深入,股价时序数据的混沌特征日益受到关注,学者们开始借助GRU进行股价预测。与LSTM网络相比,GRU通过减少门控机制来简化模型结构,但不牺牲捕捉股价动态变化的复杂模式能力。Radojičić等^[21]应用傅里叶变换提取股价新特征,并在GRU上对新

指标的性能进行基准测试,证明获得了显著的性能改进。Lv等^[22]将多个残差GRU堆叠在一起,以增加网络的深度并提高泛化能力,利用高频限价订单簿预测股价趋势,证明结果有效。Lawi等^[23]通过识别股票市场的联合运动模式,将LSTM和GRU与4种神经网络块架构相结合,提出了8种新的股价预测模型,从而提高预测结果精度。Wang等^[24]提出一种结合主成分分析和改进的GRU股价预测模型,发现该模型具有更好的预测精度和更短的训练时间。闫洪举^[25]构建了结合卷积神经网络和GRU的神经网络模型对上证指数进行预测,发现预测精度高于传统深度学习和机器学习算法。刘广等^[26]借助双向门控循环单元模型预测机构投资者业绩,发现预测结果比传统机器学习方法更优。这些研究不仅验证了GRU在金融预测中的有效性,还展示了其在多个金融预测场景中的应用潜力,为金融市场分析和预测提供了新的工具。近年来,部分研究开始在GRU中引入情感、情绪或注意力等主观信息及新闻文本信息,极大提升了股价预测结果的准确性^[27-32]。

综合来看,无论在理论探索还是在实际应用方面,PSR和GRU各自为股价预测提供了重要的视角和工具。通过结合行业相关数据重构训练集,既有模型有效缓解了深度学习在股价预测中常见的过拟合问题,为后续研究奠定了坚实的基础。但目前将二者结合起来的研究还很少,本文旨在弥补此方面研究之不足。

2 方法简介

2.1 PSR简介

PSR源于数学、物理学和统计学等学科,旨在将任意给定时间点的系统状态映射到唯一的相位点^[3]。PSR已广泛应用于多种复杂系统,如气象学中的风速和降水量、医学中的大脑活动等。

在相空间中,混沌系统的动态通常会趋向于一种特定的稳定轨迹,称为混沌吸引子。这意味着尽管混沌系统的行为看似随机,但仍存在某种有迹可循的规律。系统中的各个变量相互关联,共同推动系统向预定轨迹演进。通过分析单一变量的时间序列数据,可以揭示并还原整个系统的

底层动态规律,从而映射出系统的稳态轨道和全局动态行为^[33]。

PSR的核心在于能够在高维空间中保留原始时间序列数据的关键动力学属性和几何结构,即重构后的相空间不仅维持了原系统的结构特性,而且保留了相同的动力行为。这确保了时间序列数据的动力学特征在转换过程中得以完整保存,从而使模型能够更准确地捕捉系统的动态变化,对理解系统的长期行为及进行精确预测至关重要。

在PSR的过程中,延迟时间 τ 和嵌入维度 m 的选取至关重要,决定了相空间重构的准确性。如果 τ 过大,会导致时间序列轨迹在相空间内折叠,进而丢失重要信息;如果 τ 过小,则无法充分揭示原有吸引子的特征。同样地,如果 m 值设定过高,可能引入噪声,并且增加消耗不必要的计算资源;如果 m 值设定过低,可能产生吸引子的自交叉与重叠现象,会误导对动力学特征的识别。

目前,学界关于如何确定 τ 和 m 存在两种主要观点。一种观点认为,这两个参数应分别独立确定,比如先利用诸如自相关函数、互信息或复自相关等方法来确定 τ ,再应用几何不变量、虚假最近邻点法(False Nearest Neighbors, FNN)等来寻找适当的 m ^[34-35];另一种观点认为,这两个参数是相互关联的,因此,应利用诸如C-C方法同时确定 τ 和 m ^[36]。目前,第一种观点更受认可,本文也据此确定这两个关键参数:①借助Rosenstein等^[34]的方法,使用自相关函数来确定 τ ;②借助Kennel等^[35]的方法,使用FNN来确定 m 。

确定了 τ 和 m 后,即可借助坐标延迟重构法,将时间序列 $\{x_t\} (t=1, 2, \dots, N)$ 重构为一个 m 维的相空间 \mathbf{R}^m ,其对应的集合为 $\{\mathbf{X}_t\}$ 。二者的关系见式(1)。

$$\begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_M \end{bmatrix} = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(m-1)\tau} \\ x_2 & & x_{2+\tau} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ x_M & x_{M+\tau} & \cdots & x_{M+(m-1)\tau} \end{bmatrix}, \quad (1)$$

其中, M 为重构的相空间轨迹的相点序列长度, $M=N-(m-1)\tau$ 。

2.2 混沌性识别

混沌理论表明,在某些复杂系统中,尽管初始条件已知,结果却无法精确预测,表现出一种类似

随机的行为,即所谓的混沌现象。在这类系统中,初始条件的微小变化可能导致结果的巨大改变,即产生“蝴蝶效应”。

混沌与秩序实际上是并行存在的。混沌理论的价值在于提供了在表面上看似混乱无序的现象中寻找内在规律的可能性,揭示了即便在复杂的动态行为背后也可能存在着简单的生成机制。识别混沌现象主要有定性和定量两种策略。定性策略侧重于通过观察混沌信号在频域或时域中的独特空间结构或时间特征来进行判定,从而将其与周期性信号或其他类型的随机信号区别开来。这种策略通过直观的图形分析(如相图、庞加莱截面或时序图)识别混沌信号中的非常规模式,有助于深入理解信号的动态特性。定量策略则致力于量化描述混沌信号中的奇异吸引子,并通过计算定义明确的量化指标来判别混沌性,典型的技术包括计算最大李雅普诺夫(Lyapunov)指数、柯尔莫戈洛夫(Kolmogorov)熵或关联维数等。其中,最大 Lyapunov 指数提供了一种相对简单且普遍适用的计算框架,如果其值为正,表明系统在相空间中的轨迹将以指数速度发散,则被认为具有混沌属性。本文也采用它来识别混沌性。

沃尔夫(Wolf)方法被广泛应用于计算最大 Lyapunov 指数,以确定系统是否具有混沌特性^[37]。这种方法通过分析系统轨迹在相空间中的发散速度来评估初始条件的微小变化是如何影响系统行为演变的。Wolf 方法具有普遍适用性和精确性,特别在数据量不是非常大的情况下,其稳定性和可靠性更佳。本文也使用该方法。

2.3 循环神经网络简介

RNN 是一种应对非线性时间序列数据分析而设计的神经网络。与传统神经网络模型相比,RNN 考虑了数据之间的时序关系,因此有更为优良的表现。

RNN 的架构主要包括输入层、一个或多个循环隐藏层以及输出层。在每一个时间步,隐藏层不仅接收当前时间步的输入数据,还会接收来自前一时间步的隐藏层状态。基于这两个输入,隐藏层会更新其当前状态,并将这一状态传递至下一时间步的隐藏层及输出层,以此实现数据在时间序列中的传递和学习。这种结构使得 RNN 特别适合处理和预测与时间序列相关的数据,因为它能够

记忆之前的信息并影响后续输出。RNN 允许使用固定的参数集在整个序列处理过程中持续运作,意味着网络参数的数量不会因序列长度的延伸而增加,从而在模型设计上实现了高效的参数使用。

标准 RNN 在处理长序列数据时可能面临梯度消失或梯度爆炸的问题,这限制了它在学习长距离依赖关系时的效能。为了克服这些问题,发展了更高级的 RNN 结构,如 LSTM 和 GRU,这些结构通过特殊的门控机制来调节信息流,有效地保存长期依赖信息,并在必要时忘记无用信息。

GRU 与 RNN 最主要的区别在于 GRU 对隐状态采用了门控机制。这种机制能够控制信息在单元内的流动,帮助 GRU 更好地在不同的时间步骤中维持关键信息并舍弃不重要的信息。GRU 通过两种特殊的门机制来管理信息流,即重置门和更新门。重置门旨在帮助模型捕捉时间序列中的短期依赖性,而更新门则用于维持长期依赖的信息。这种结构使 GRU 能够在不同时间点灵活调整其内部状态,优化对时间序列数据的处理效果,提高预测准确性和稳定性。

3 研究设计

3.1 样本与平台选择

参考 Zhang 等^[8]、Niu 等^[27]和段虎等^[11-12]的做法,选择 5 个代表性股指(上证指数、深圳成指、深圳综指、沪深 300 指数和中证 1000 指数)及 5 只大市值的权重股(贵州茅台、招商银行、工业富联、比亚迪和科大讯飞)作为样本。样本期选择 2020—2023 年共 970 个交易日。指数和个股收盘价数据从 Wind 数据库下载。

股价预测模型采用 Python 语言编写,开发环境为 Visual Studio Code。该预测模型基于 TensorFlow 深度学习框架构建,并在搭载 Windows 10 操作系统的计算机上,利用 TensorFlow 的 GPU 支持版本进行训练,而 PSR 部分则借助 Octave 软件平台完成。

3.2 PSR-GRU 模型预测步骤

(1)使用 PSR 对原始输入数据进行深加工,将一维时间序列映射成 m 维序列。重构得到的多维序列相比原始一维序列更加准确地反映了数据内在特性和相互关系。

(2)采用 Wolf 方法对序列的最大 Lyapunov 指数进行估算,据此评估序列的混沌特性。当最大 Lyapunov 指数为正时,序列被认定为具有混沌特性,使用 PSR 处理是合适的;若最大 Lyapunov 指数为负,序列未显示混沌特性,则不采用 PSR,并且不进一步探讨这些非混沌序列的预测结果,终止实验。

(3)股价预测。经过 PSR 处理后的数据被用作 GRU 网络的输入以预测股票价格。①对经过 PSR 后的数据进行零均值归一化,并使用滑动窗口方法将其转换为时间步长为 30 的时间序列数据,以保持趋势信息的完整性,并以 0.8 的比例划分训练集与测试集;②合理设计 GRU 网络架构,确定模型的超参数如隐藏层的神经元数量、随机失活(Dropout)率、序列长度、循环训练次数以及循环训练的小批量样本数(Batch)。考虑到股票价格预测的回归本质,选择线性激活函数作为输出层。模型经过充分训练后,使用测试数据进行评估以获得预测值,并对这些预测值进行逆标准化处理,以获得最终的股价预测结果。

(4)评估预测结果。评估方法主要有两种:①借助曲线图将模型生成的预测值与实际观测值进行对比。如果两种曲线吻合度较高,表明预测结果较准确。②观察损失函数随训练迭代次数的变化。在理想情况下,训练集和测试集上的损失值会随着迭代次数的增加而逐渐降低,这表明模型正在学习并改进其预测能力。如果仅训练集上的损失值下降而测试集上的损失值上升,则可能表示模型过度适应了训练数据,即发生了过拟合,这会降低模型在新数据上的预测准确性;反之,如果训练集上的损失值升高,这通常意味着模型未能从迭代中学习得到有效信息,导致性能退化;如果损失值在迭代过程中无显著下降趋势,且保持相对平稳,则可能表明模型未能捕捉到数据中的规律,需要进一步优化模型结构或参数。

3.3 参数设置

在构建 GRU 预测模型时,主要使用 Tensorflow 框架中的 Keras 库。模型的核心组成包括层(Layers)、损失函数(Losses)和优化器(Optimizers)3个部分。在 Layers 中,主要调用了 GRU、Dense、Activation 和 Dropout 等模块。

本文使用的模型 Layers 包括 GRU 和 Dense

层。由于数据量有限且并不复杂,所使用的模型共搭建了 5 层网络结构,包括 4 层带有内置 Dropout 正则化的 GRU 层和 1 层全连接的 Dense 层。这种设计通过在每个 GRU 层内部应用 Dropout 来防止过拟合,增强模型的泛化能力,而最终的 Dense 层则负责输出预测结果。

模型训练的优化器使用 Adam 优化器。作为一种效率较高的一阶优化方法,Adam 优化器相较于传统的随机梯度下降算法表现更为优异,因其简明的实现、高效的计算性能及快速达到较优结果的能力而被广泛采用。

模型参数主要包括循环训练的小批量样本数(Batch)、迭代次数(Epoch 和 Dropout)、训练集与测试集分割比例(Train_Test_Split)、GRU 神经元个数(GRU_Units)及序列步长(Sequence_Length)等。其中,选择恰当的序列步长对预测准确性至关重要。序列步长决定了在进行预测时,将多少个历史数据作为输入,既不应过短,也不宜过长。本文使用的是日数据,通过对比分析,将序列步长设定为 1,即仅采用上一期的信息进行预测时,模型更加聚焦于最近的数据点,从而能够达到更好的预测效果,且更易于为普通投资者所利用。最终参数设置见表 1。

表 1 模型参数
Table 1 Model parameters

| 参数 | 取值 |
|------------------|-------|
| Batch Size | 16.0 |
| Epoch | 100.0 |
| Dropout | 0.2 |
| Train_Test_Split | 0.8 |
| GRU_Units | 256.0 |
| Sequence_Length | 1.0 |

4 股价预测分析

4.1 序列的相空间重构

本文分别利用自相关函数法(Autocorrelation Function, ACF)和 FNN 确定延迟时间及嵌入维数,对 5 个主要股指及 5 只权重股分别进行计算。以下为部分计算结果展示,以上证指数 000001.SH 为例。

利用 ACF 计算延迟时间的判断标准是,当

ACF 值首次穿越零界限,即从正相关性降至无相关性时,这个特定的滞后点被定义为延迟时间。这一点标志着序列中早期值对后续值的直接影响已经减弱到统计上不显著的程度。据此可测算得到上证指数 000001.SH 对应的延迟时间为 174 (图 1)。

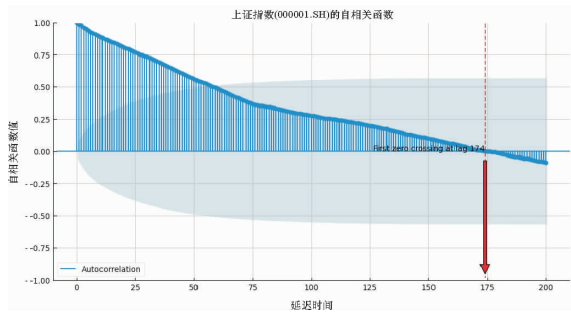


图 1 上证指数的延迟时间

Fig. 1 τ of the Shanghai composite index

在利用 FNN 方法计算最佳嵌入维度时,遵循的判断标准是基于时间序列数据重构相空间时,假邻居的比例急剧下降并最终趋于平稳。在嵌入维度较低时,由于维度的限制,可能会出现高量的假邻居;随着嵌入维度的增加,相空间得以展开,假邻居数量急剧下降。理想的嵌入维度应在 FNN 比率首次下降到一个较低水平且趋于稳定的点确定。如图 2 所示,嵌入维度从 1.00 增加到 2.00 时,FNN 比率显著下降,随后从 2.00 增加到 3.00 时,比率进一步降低。然而,在抵达维度 3.00 时,FNN 不再下降,表明增加更多的维度并没有显著减少假邻居的数量。因此,确定最佳嵌入维度为 3.00。

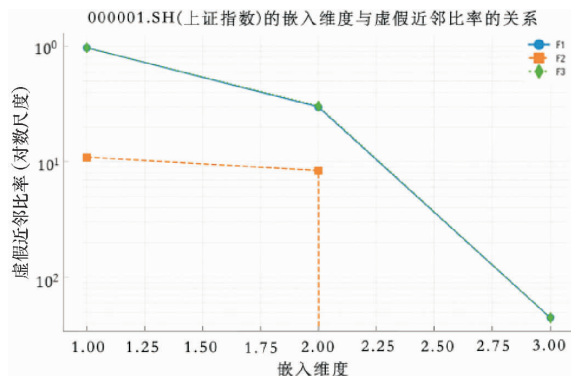


图 2 上证指数最佳嵌入维度

Fig. 2 m of the Shanghai composite index

将测算得到的 5 个股指及 5 只个股的 τ 和 m 进行汇总,结果见表 2。

表 2 参数结果

Table 2 Parameters results

| 实验样本 | Obs | τ | m | M | λ |
|---------|-----|--------|------|-----|-----------|
| 上证指数 | 970 | 174 | 3.00 | 622 | 0.031 8 |
| 深证成指 | 970 | 249 | 3.00 | 472 | 0.034 9 |
| 深证综指 | 970 | 181 | 3.00 | 608 | 0.028 5 |
| 沪深 300 | 970 | 247 | 3.00 | 476 | 0.033 7 |
| 中证 1000 | 970 | 100 | 3.00 | 770 | 0.025 3 |
| 贵州茅台 | 970 | 164 | 3.00 | 642 | 0.024 1 |
| 招商银行 | 970 | 187 | 3.00 | 596 | 0.034 0 |
| 工业富联 | 970 | 135 | 3.00 | 700 | 0.036 1 |
| 比亚迪 | 970 | 279 | 3.00 | 412 | 0.036 6 |
| 科大讯飞 | 970 | 117 | 3.00 | 736 | 0.028 3 |

从表 2 可以看到各个指数和个股在 PSR 的过程中所确定的延迟时间和 τ 嵌入维度 m 。整体而言,指数和个股的嵌入维度均为 3.00,这表明在三维空间中,这些时间序列的动态特性得到了充分展现,而更高维度的增加对于揭示系统的动力学结构并无显著贡献。另外,尽管选定的指数和个股在嵌入维度上显示出一致性,延迟时间的差异揭示了不同指数及个股在市场反应速度和价格变动记忆长度上的个体差异。在 PSR 的后续步骤中,应当根据各自的参数特征来定制重构过程,从而得到各具特点的重构序列。

完成 PSR 的相关参数设定后,便可以对各样本的收盘价进行处理,通过将单维时间序列转换为多维形式,形成新的重构时间序列。仍以上证指数为例,由于上证指数在样本期内共有 970 个交易日的收盘价数据,根据重构公式 $M = N - (m - 1) \times \tau$,得到的重构时间序列长度是 622(表 2)。

4.2 序列的混沌性识别

在本研究中,为了辨识时间序列的混沌特性,采用最大 Lyapunov 指数作为评估标准。最大 Lyapunov 指数是一种衡量系统对初始条件变化反应灵敏度的量度。根据判断准则,如果该指数值大于零,则表明时间序列展现出混沌行为,意味着在相空间里,接近的轨迹会以指数速度发散;相反,若最大 Lyapunov 指数小于零,则序列被认为不具备混沌特性,即接近的轨迹会最终趋于合并。

对选取的 5 个指数和 5 只个股的最大 Lyapunov 指数进行计算,发现所有样本的指数值均为正,意味着样本的日收盘价序列均展现出混沌特征,是混沌时间序列。这验证了关于 PSR 可应用于股价预测中的构想。此外,也说明股价序列对

初始条件表现出高度敏感性,即最初极小的变化能随时间发展而放大,导致预测的准确性随着时间的推移而下降。因此,长期预测的结果是不可靠的,只能进行短期预测。

4.3 模型性能检验

经过反复试验,本文确定了网络架构及参数,将 GRU 运用于单步股价预测。以上证指数和贵州茅台为例,图 3 和图 4 展示了它们各自的预测价格与实际股价之间的比较。

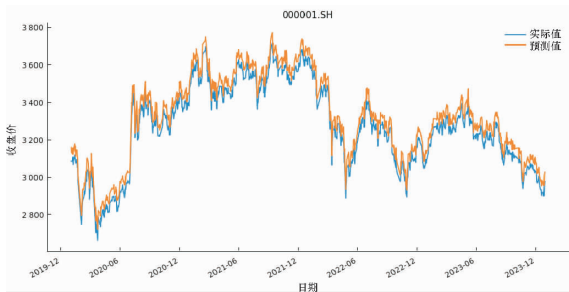


图 3 上证指数预测结果对比

Fig. 3 Predicted and actual prices of Shanghai composite index

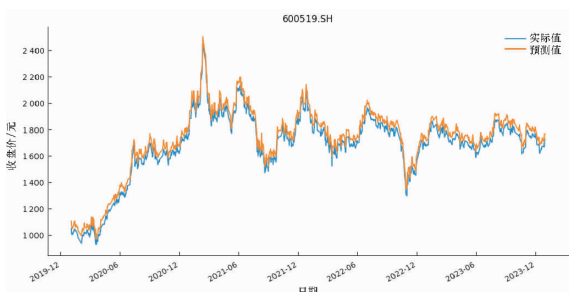


图 4 贵州茅台预测结果对比

Fig. 4 Predicted and actual prices of Kweichow Moutai Co., Ltd.

观察图 3 和图 4 可发现,模型产生的预测值与实际值高度一致。该现象在所有测试样本中均有观测到,这证明所采用的网络结构在进行单步预测方面表现出色。图 5 和图 6 为两个样本单步预测的损失函数图。

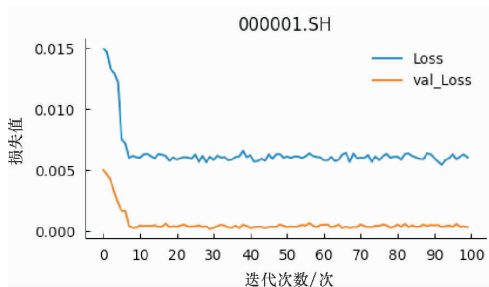


图 5 上证指数 Loss 函数

Fig. 5 Loss function of Shanghai composite index

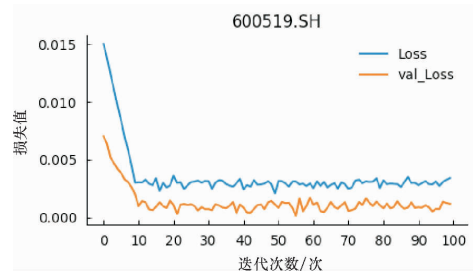


图 6 贵州茅台 Loss 函数

Fig. 6 Loss function of of Kweichow Moutai Co., Ltd.

由图 5 和图 6 可知,两个样本在训练集 (Loss) 和测试集 (Val_loss) 上的损失函数都呈现下降趋势,表明模型训练过程中没有发生过拟合现象,训练效果是理想的。同时,其他样本的实验数据也得到类似的下降趋势,进一步证实了所使用的网络结构的有效性。

损失函数通常分为分类和回归两大类。借鉴刘广等^[32]的研究,使用均方误差 (Mean Squared Error, MSE) 和平均绝对误差 (Mean Absolute Error, MAE) 作为损失函数,其中, MSE 通过计算预测值与实际值之间差异的平方和的平均值来评估预测模型的准确性, MAE 评估预测模型的整体性能。一般来说, MSE 和 MAE 值越小,模型预测精度越高,性能越好。计算方法分别见式(2)和式(3)。

$$MSE = \frac{1}{N} \sum_{t=1}^N (\hat{y}_t - y_t)^2, \quad (2)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |\hat{y}_t - y_t|, \quad (3)$$

其中, \hat{y}_t 和 y_t 分别代表 t 时刻的预测值和实际值, N 代表测试集中测试数据总数。

损失值结果见表 3。损失值总体都较小,表明预测模型的精度较高且性能良好。

表 3 损失值结果

Table 3 Results of loss function

| 样本 | 训练集 | | 测试集 | |
|---------|---------|---------|---------|---------|
| | MSE | MAE | MSE | MAE |
| 上证指数 | 0.006 5 | 0.003 8 | 0.002 1 | 0.001 5 |
| 深证综指 | 0.004 2 | 0.002 7 | 0.002 3 | 0.001 2 |
| 深证成指 | 0.003 7 | 0.001 5 | 0.002 1 | 0.001 1 |
| 沪深 300 | 0.005 6 | 0.005 2 | 0.003 4 | 0.002 6 |
| 中证 1000 | 0.004 5 | 0.003 6 | 0.002 9 | 0.002 1 |
| 贵州茅台 | 0.004 3 | 0.002 9 | 0.001 9 | 0.001 0 |
| 招商银行 | 0.029 3 | 0.002 4 | 0.029 3 | 0.016 3 |
| 工业富联 | 0.003 7 | 0.003 1 | 0.002 0 | 0.000 9 |
| 比亚迪 | 0.003 2 | 0.002 8 | 0.001 4 | 0.000 6 |
| 科大讯飞 | 0.004 5 | 0.002 1 | 0.003 1 | 0.002 4 |

4.4 预测结果对比分析

仍以上证指数 000001.SH 为例,图 7 和图 8 分别展示了基于 LSTM 和 SVM 的预测结果。

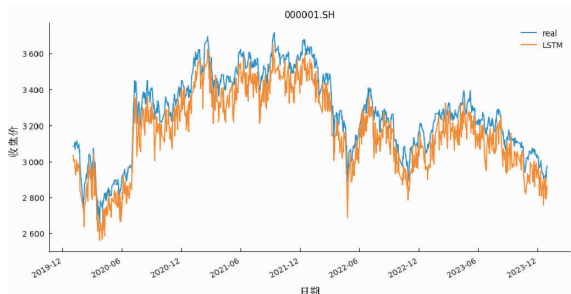


图 7 LSTM 对上证指数的预测结果

Fig. 7 Prediction results by LSTM of Shanghai composite index

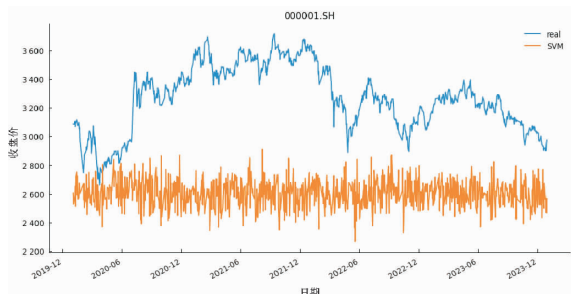


图 8 SVM 对上证指数的预测结果

Fig. 8 Prediction results by SVM of Shanghai composite index

对比图 3、图 7 和图 8 可以发现,PSR-GRU 模型在股价预测中展示出较好的性能,且预测精度更高,更紧密地预测了股价走势,证明其在股价预测

中的可靠性。与之相比,SVM 在预测精度方面的表现则不尽人意,预测价格与真实价格之间的差异较大,且在股价剧烈波动期间,这种差异尤为显著。

5 结 论

本文结合 PSR 和 GRU 对股价走势进行预测。在 A 股中选取 5 个主要指数和 5 只市值较大的权重股为样本,进行以下研究:①借助 ACF 和 FNN 确定延迟时间和嵌入维度,并使用 PSR 对股价时间序列进行相空间重构;②借助 Wolf 方法计算最大 Lyapunov 指数,揭示股价波动的混沌特征;③使用 GRU 对股价进行预测,并将预测结果与经典预测模型 LSTM 及 SVM 所得预测结果进行对比。结果显示,股价序列具有混沌性,PSR-GRU 模型相较于传统的 LSTM 和 SVM 具有更好的预测性能。

本文研究仍存在有待完善的地方,比如暂时仅实施了单步预测,未来可扩展至多步预测,以探究模型在更长时间尺度上的预测性能。此外,暂时仅采用基于单一变量(股价)的时间序列数据进行相空间重构,使用的信息可能不足。实际上,多变量时间序列能够提供更为丰富且综合的市场信息,未来可考虑融合诸如成交量、换手率、市盈率等指标,或纳入情绪、新闻文本等信息进行综合分析,不断提高预测精度。

参考文献:

- [1] Lévy P. Théorie de l'addition des variables aléatoires[M]. Paris: Gauthier-Vilars, 1937.
- [2] Peters E E. Chaos and order in the capital markets: a new view of cycles, prices, and market volatility[M]. 2nd ed. New York: Wiley, 1996.
- [3] Packard N H, Crutchfield J P, Farmer J D, et al. Geometry from a time series[J]. Physical Review Letters, 1980, 45(9):712-716.
- [4] Lang S H, Zhu H, Sun G D, et al. A study on methods for determining phase space reconstruction parameters[J]. Journal of Computational and Nonlinear Dynamics, 2022, 17(1):011006.
- [5] Dousseh P Y, Ainamon C, Miwadinou C H, et al. Chaos control and synchronization of a new chaotic financial system with integer and fractional order[J]. Journal of Nonlinear Sciences and Applications, 2021, 14(6):372-389.
- [6] Huang Y, Wan J S. Hierarchical analysis of Chinese financial market based on manifold structure[J]. Annals of Operations Research, 2022, 315(2):1135-1150.
- [7] Shah Z, Bonyah E, Alzahrani E, et al. Chaotic phenomena and oscillations in dynamical behaviour of financial system via fractional calculus[J]. Complexity, 2022, 2022(1):8113760.
- [8] Zhang H G, Liang J Z, Chai Z L. Stock prediction based on phase space reconstruction and echo state networks[J]. Journal of Algorithms & Computational Technology, 2013, 7(1):87-100.
- [9] Khuwaja P, Khowaja S A, Khoso I, et al. Prediction of stock movement using phase space reconstruction and extreme learning machines[J]. Journal of Experimental & Theoretical Artificial Intelligence, 2020, 32(1):59-79.

- [10] Li C T, Liu K. Path signature-based phase space reconstruction for stock trend prediction[J]. *International Journal of Data Science and Analytics*, 2022, 14(3):293-304.
- [11] 段虎,沈菲. 上证指数中的混沌现象研究[J]. *数量经济技术经济研究*, 2002, 19(2):91-94.
- [12] 段虎,胡斌. 深成指数的混沌吸引子分形维研究及预测[J]. *数量经济技术经济研究*, 2002, 19(4):107-109.
- [13] 王卫宁,汪秉宏,史晓平. 股票价格波动的混沌行为分析[J]. *数量经济技术经济研究*, 2004, 21(4):141-147.
- [14] 谢朝华,李忠,郑咏梅,等. 中国股票市场分形与混沌特征:1994—2008[J]. *系统工程*, 2010, 28(6):30-35.
- [15] 陈佐,谢赤,李晓东. 基于相空间重构技术的EM聚类模糊神经网络预测模型及其应用[J]. *中国软科学*, 2006(8):147-153.
- [16] 郭兆纲,胡振鹏. 一种测度金融历史事件相关性的新方法[J]. *数量经济技术经济研究*, 2015, 32(9):101-118.
- [17] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786):504-507.
- [18] 苏治,卢曼,李德轩. 深度学习的金融实证应用:动态、贡献与展望[J]. *金融研究*, 2017(5):111-126.
- [19] Jin Z B, Jin Y X, Chen Z Y. Empirical mode decomposition using deep learning model for financial market forecasting[J]. *PeerJ Computer Science*, 2022, 8: e1076.
- [20] 马甜,姜富伟,唐国豪. 深度学习与中国股票市场因子投资——基于生成式对抗网络方法[J]. *经济学(季刊)*, 2022, 22(3):819-842.
- [21] Radojičić D, Kredatus S. The impact of stock market price Fourier transform analysis on the Gated Recurrent Unit classifier model[J]. *Expert Systems with Applications*, 2020, 159: 113565.
- [22] Lv X, Zhang L. *International Conference on Knowledge Science, Engineering and Management*, August 14-16, 2021[C]. Berlin: Springer, 2021.
- [23] Lawi A, Mesra H, Amir S. Implementation of long short-term memory and gated recurrent units on grouped time-series data to predict stock prices accurately[J]. *Journal of Big Data*, 2022, 9(1):89.
- [24] Wang J, Liu D, Jin L, et al. A PCA-IGRU model for stock price prediction[J]. *Journal of Internet Technology*, 2023, 24(3):621-629.
- [25] 闫洪举. 基于深度学习的金融时间序列数据集集成预测[J]. *统计与信息论坛*, 2020, 35(4):33-41.
- [26] 刘广,刘汉中. 多约束条件放松下机构投资者资产组合变动与业绩检验和预测[J]. *数理统计与管理*, 2024, 43(6):1104-1120.
- [27] Niu H L, Xu K L. A hybrid model combining variational mode decomposition and an attention-GRU network for stock price index forecasting[J]. *Mathematical Biosciences and Engineering*, 2020, 17(6):7151-7166.
- [28] Xu H F, Chai L, Luo Z M, et al. Stock movement prediction via gated recurrent unit network based on reinforcement learning with incorporated attention mechanisms[J]. *Neurocomputing*, 2022, 467: 214-228.
- [29] Haryono A T, Sarno R, Sungkono K R. Transformer-gated recurrent unit method for predicting stock price based on news sentiments and technical indicators[J]. *IEEE Access*, 2023, 11: 77132-77146.
- [30] 谷丽琼,吴运杰,逢金辉. 基于 Attention 机制的 GRU 股票预测模型[J]. *系统工程*, 2020, 38(5):134-140.
- [31] 崔婷,黄斐然. 基于情感分析大模型的股票预测:结合 GRU 和 ALBERT 的预测模型[J]. *东岳论丛*, 2024, 45(2):113-123.
- [32] 刘广,易鸿. 融合媒体信息和信号分解的股票市场深度学习预测[J]. *计算机科学*, 2024, 51(S1):1104-1115.
- [33] Takens F. *Dynamical Systems and Turbulence*, Warwick 1980[C]. Berlin:Springer,1981.
- [34] Rosenstein M T, Collins J J, De Luca C J. Reconstruction expansion as a geometry-based framework for choosing proper delay times[J]. *Physica D: Nonlinear Phenomena*, 1994, 73(1/2):82-98.
- [35] Kennel M B, Brown R, Abarbanel H D. Determining embedding dimension for phase-space reconstruction using a geometrical construction[J]. *Physical Review A: Atomic, Molecular, and Optical Physics*, 1992, 45(6):3403-3411.
- [36] Kim H S, Eykholt R, Salas J D. Nonlinear dynamics, delay times, and embedding windows[J]. *Physica D: Nonlinear Phenomena*, 1999, 127(1/2):48-60.
- [37] Wolf A, Swift J B, Swinney H L, et al. Determining Lyapunov exponents from a time series[J]. *Physica D: Nonlinear Phenomena*, 1985, 16(3):285-317.