

# 暨南大学硕士学位论文

题名（中英对照）：BP 和 RBF 网络模型在黄金价格预测中的应用

Application of BP and RBF Network Models in Gold Price Forecast

作者姓名：乔宏图

指导教师姓名  
及学位、职称：范旭乾  
副教授

学科、专业名称：应用统计

学位类型：专业学位

论文提交日期：2015 年 6 月

论文答辩日期：2015 年 5 月 26 日

答辩委员会主席：田联房

论文评阅人：盲审

学位授予单位和日期：2015 年 6 月 26 日

# 独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得暨南大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：

签字日期： 2016 年 6 月 日

## 摘要

本文主要是通过 BP 神经网络和 RBF 神经网络对黄金价格进行预测，并比较两种模型预测效果。

首先用功率谱法和李雅普诺夫指数对 2015 年黄金交易价格的数据进行分析，得出黄金价格序列是非线性的结论。然后将黄金价格时间序列进行空间重构，计算最小嵌入维数，并把最小嵌入维数作为神经元节点的个数。最后用 BP 和 RBF 神经网络对黄金交易价格的数据进行模拟和预测。通过比较两种模型对黄金价格的预测结果，发现两种模型都可以用来对黄金价格进行预测，并且 RBF 神经网络模型在黄金价格预测方面比 BP 神经网络模型更有效。

关键字：黄金价格，空间重构，BP 神经网络，RBF 神经网络

## **Abstract**

This thesis mainly uses the BP neural network and RBF neural network to forecast the gold price, and compares the performance of two kinds of models for predicting the gold price.

First of all, we will analyze the gold price data in 2105 by power spectrum method and Lyapunov index. We draw a conclusion that gold price time series is nonlinear. Then we will reconstruct the space of the gold price and calculate the minimum embedding dimension. We will use the minimum embedding dimension as the number of neuron nodes. Finally, we will use BP and RBF neural network to simulate and forecast the data of gold price. By comparing the two kinds of forecast results, we find that both models can well predict the gold price, and we also find that RBF neural network model in the gold price forecast is more effective than BP neural network model.

**Keywords:** Gold price, Space reconstruction, BP neural network, RBF neural network

## 目录

摘要.....	I
Abstract.....	II
第一章、绪论.....	1
1.1 研究背景及意义.....	1
1.2 文献综述.....	2
1.3 本文研究内容.....	3
1.4 本文的创新点.....	3
1.5 本文结构.....	4
第二章、黄金市场介绍.....	5
2.1 黄金市场发展历史.....	5
2.2 国际主要黄金市场.....	5
2.3 黄金价格的影响因素.....	6
2.4 本章小结.....	8
第三章、BP 和 RBF 神经网络模型简介.....	9
3.1 神经网络模型发展过程.....	9
3.2 神经网络的特点和功能.....	10
3.3 神经网络常见的连接形式.....	11
3.4 神经网络的学习方法.....	12
3.5 BP 神经网络.....	13
3.6 RBF 神经网络.....	17
3.7 本章小结.....	21
第四章、时间序列的非线性判定.....	22
4.1 混沌的定义、特征及判别方法.....	22
4.2 定性分析.....	23
4.3 定量分析.....	25
4.4 本章小结.....	26
第五章、时间序列的空间重构.....	27
5.1 Takens 定理.....	27
5.2 确定延迟时间.....	28

---

5.3 确定嵌入维数.....	29
5.4 本章小结.....	31
第六章、神经网络模型预测黄金价格.....	32
6.1 数据的预处理.....	32
6.2 相关性分析.....	33
6.3 BP 神经网络预测黄金价格.....	34
6.5 BP 神经网络和 RBF 神经网络预测结果比较.....	43
6.6 本章小结.....	44
第七章、总结与展望.....	45
参考文献.....	46
致谢.....	49

## 第一章、绪论

### 1.1 研究背景及意义

黄金自十九世纪被大量开采以来,一直以其独有的贵金属地位在国际金融市场中占据着重要位置。二战结束后形成的布雷顿森林体系使得黄金成为和美元挂钩的贵金属,黄金储备直接可以决定一个国家的货币币值。后来布雷顿森林体系瓦解,黄金不再是世界上最重要的货币储备,但仍然是各国重要战略储备之一。黄金的重要作用主要体现在国家黄金储备、居民消费市场和金融投资等方面。在国家黄金储备方面,我国作为一个新兴的大国,正在不断通过国际黄金市场增加黄金储备量,截止 2016 年一季度末,我国黄金储备量已经达到 1797.5 吨。在居民消费市场方面,因为黄金是首饰配饰的重要原料之一,因此黄金价格的变化对每一名消费者都有很大的影响。在金融投资方面,因为黄金价格的波动相对于中国股票市场的波动更小,价格更加稳定,因此在经济动荡的时候,可以通过投资黄金产品来有效规避市场价格波动风险。鉴于黄金的重要作用,研究黄金价格的波动是有意义的。通过了解以前学者对黄金价格的研究,发现黄金是一种受多重因素影响的<sup>[1-4]</sup>金融工具。因此本文将综合多种因素对黄金价格进行预测。

黄金价格在上世纪 40 年代只有 35 美元/盎司(折合人民币即期汇率 7.29 元/克),在 2011 年 9 月 6 日达到 1920.8 美元/盎司(折合人民币 400.04 元/克)的历史最高点。过去七十多年的时间里,黄金价格翻倍近 55 倍。就 2015 年而言,中国黄金市场的黄金价格就经历了从 260.8 元/克到 216.9 元/克的波动,波动幅度达到了 43.9 元/克,占到最低价位时价格的 20.24%。黄金价格的波动对黄金消费市场和黄金的生产都产生重大影响。在此举例说明黄金价格波动对黄金消费市场的影响:黄金饰品市场上的普通黄金项链的质量一般在 20-50 克之间,不考虑工艺等影响黄金饰品价格的其他因素,只考虑黄金成本,43.9 元的价差对 50 克黄金项链的价格影响为 2195 元,占到国家统计局公布的 2015 年人均可支配收入 21966 元的 9.99%。对整个中国黄金消费市场,在 2015 年黄金消费量被印度超越之前,我国一直是黄金消费世界第一大国,2015 年黄金消费量达到 985.9 吨<sup>[5]</sup>。近千吨的消费量意味着黄金价格的轻微波动,对整个消费市场的价值总量影响都是巨大的。在黄金的生产方面,2011 年我国黄金产量达到 361 吨,与 2005 年相比,年均增长 8.3%。2012 年中国黄金产量突破 400 吨大关,达到 403.1 吨。2013 年,中国黄金产量达到 428.163 吨。2014 年,中国黄金产量达到 451.799 吨<sup>[6]</sup>。2015 年,中国黄金产量达到 450.053 吨,黄金产量虽然首次下跌,但仍连续九年位居世界第一。作为世界最大的黄金生产国,黄金价格的波动,

对整个市场上黄金生产企业总利润的影响也是巨大的。

通过模型合理的预测黄金价格变化，可以更好地趋利避害，帮助黄金市场的参与者选择对自己最有利的方式来进行黄金交易。对黄金市场消费者而言，可以选择价格较低的时候进行黄金消费，降低消费成本；对黄金市场生产企业而言，可以通过合理规划产量实现利润最优；对国家而言，可以在黄金价格低位增持黄金储备，降低储备黄金的成本；对于众多投资者而言，可以通过价格预测抓住市场交易时机，在投资市场获取投资收益。综上所述，黄金价格预测研究具有重要意义。

## 1.2 文献综述

### 1.2.1 国外文献综述

黄金一直在世界经济体系中扮演者重要的作用，外国对相关研究也比较全面。在黄金和通货膨胀关系方面：1977年，Jastram 和 Roy W 分析了黄金价格、美国消费指数等相关因素，得到黄金是对抗通货膨胀有力工具的结论<sup>[7]</sup>。2013年，J Beckmann 和 R Czudaj 通过运用向量误差修正模型对1970年到2011年进行分析，分析了黄金在不动地区和不同时间范围内对冲通货膨胀的效果<sup>[8]</sup>。2014年，JA Batten, C Ciner 和 BM Lucey 通过运用 VAR 模型和协整检验对1985年以后黄金价格和 CPI 的分析，得出黄金有对抗通货膨胀的作用，说明黄金的商品货币属性<sup>[9]</sup>。在黄金价格和美元及原油价格方面：2010年，PK Narayan, S Narayan, X Zheng 研究了黄金价格和石油期货关系，表明可以通过一个市场价格来预测另一个市场的价格<sup>[10]</sup>。2013年，YS Wang 和 YL Chueh 通过分析，得出黄金价格、原油价格和美元利率及美元之间是有相互影响的<sup>[11]</sup>。2013年，HF Chang, LC Huang 和 MC Chin 通过运用协整检验、VAR 模型、脉冲响应分析、格兰杰因果检验和方差分解方法分析黄金价格、石油价格及汇率三者之间的关系，得到三者在很大程度上是相互独立的结论<sup>[12]</sup>。还有学者对黄金期货价格进行分析。比如：2006年，G Grudnitski, L Osburn 运用神经网络模型对1983年到1990年的黄金期货和标准普尔期货以月数据为单位进行预测分析<sup>[13]</sup>。同时也有学者通过不同的模型预测黄金价格。比如：2009年，Z Ismail, A Yahya 和 A Shabri 运用多元线性回归模型，引入大宗商品、汇率、通货膨胀率等指标对黄金价格进行预测<sup>[14]</sup>；2014年，A Sharma 和 S Baby 通过 ARIMA 时间序列模型对1990年2月到2014年9月的黄金价格进行预测分析<sup>[15]</sup>。除了学者对黄金价格有着诸多研究，国外的相关网站，也对黄金价格进行相关近期预测。例如：goldsliverworlds 网站在2014年时对2015年黄金价格整体变化进行预测<sup>[16]</sup>；forecast 网站在2016年对本年黄金价格走势进行预测<sup>[17]</sup>。

### 1.2.2 国内文献综述

在相关性分析方面：2004年，杨柳勇和史震涛通过分析1973年到2000年的数据黄金价格数据，说明黄金价格的长期决定因素有汇率、股票指数、通货膨胀率和利率四个因素<sup>[18]</sup>。2008年，刘曙光和胡再勇分时段研究黄金价格和股价指数、美国消费者价格指数、利率、黄金储量等因素，表明在不同时段影响黄金价格的因素不能完全保持一致并且黄金价格和美元的价格变动趋于一致<sup>[19]</sup>。2009年，付丹、梅雪和张晖通过菲利普斯曲线和最小二乘法研究1996年到2007年中国通货膨胀和黄金价格波动的关系，认为黄金价格不仅可以对抗通货膨胀，还可以预测经济走势<sup>[20]</sup>。2007年，杨叶通过对黄金价格和石油价格波动的比较，说明黄金和是有的价格波动大多时候呈正向联动关系<sup>[21]</sup>。在价格预测方面：1999年，胡乃联和宋鑫通过自适应过滤模型对1968年到2000年黄金价格进行预测<sup>[22]</sup>。2010年，林雨、孔刘柳和刘培运用ARFIMA模型对2002年到2009年黄金价格进行模拟预测，说明了黄金价格具有长期记忆性<sup>[23]</sup>。此外涂小龙运用GM模型预测了2011年黄金价格<sup>[24]</sup>。除了相关学者的研究，国内的一些财经网站，诸如和讯网、新浪财经、中国黄金投资网等也对黄金价格走势进行预测。

### 1.3 本文研究内容

首先对黄金价格的时间序列进行非线性判别。选取的判别方式是通过判断黄金价格时间序列是混沌的，进而说明是非线性的。采用的方法是定性判别的功率谱法和定量判别的Lyapunov指数法。

然后对黄金价格的时间序列进行空间重构。采用互信息法确定延迟时间，采用CAO方法确定最小嵌入维数，并且把最小嵌入维数作为神经网络神经元个数。

最后运用BP神经网络和RBF神经网络对黄金价格进行预测，并比较两种模型对黄金价格预测的效果。在神经网络的输入数据选取上，分析汇率、上证指数、原油价格、开盘价等因素和黄金价格的相关性，选取相关性高变量的作为参考的输入变量。在预测上，分别比较BP神经网络和RBF神经网络在10期和30期预测的误差平方和和均方误差，结果表明RBF神经网络模型预测效果更好。

### 1.4 本文的创新点

首先在数据方面，以往分析都是选取月或者周的数据进行长期分析，而本文选用的是每日的黄金价格数据进行分析。其次在神经网络的神经元节点个数选择上，选取空间重构的最小嵌入维数作为神经元个数。然后在预测模型的输入数据上，选择的不只是汇率、股

指、原油价格等其他因素，还包括自身的开盘价和成交量等自身的因素，综合考虑到神经网络模型中对结果进行预测。最后比较 BP 神经网络模型和 RBF 神经网络模型对黄金价格的预测结果。

## 1.5 本文结构

本文的具体结构如下：

第一章：绪论，介绍本文的写作背景和研究意义、国内外研究现状、本文的主要研究内容。主要说明本文写作的现实意义以及所应用的理论发展现状。

第二章、具体描述黄金市场的相关知识，主要包括黄金市场历史、国际上主要黄金市场以及黄金价格的影响因素。

第三章、详细介绍神经网络模型相关知识，重点介绍本文用来预测的 BP 和 RBF 神经网络模型。

第四章、时间序列的非线性判定，从混沌理论的定义入手，介绍混沌定义、特征、判定方法以及相关判别方法的数学原理，并分别选取定性和定量判别的方法对测试数据进行混沌判别。

第五章、空间重构。说明空间重构原理，空间重构的主要内容包括确定延迟时间  $\tau$  和嵌入维数  $m$ ，说明延迟时间和嵌入维数在空间重构过程中的作用，介绍相关空间重构的方法和数学原理，选取互信息法确定延迟时间  $\tau$ ，选取 CAO 方法确定嵌入维数  $m$ 。

第六章、神经网络预测模型。本文的较重要章节。研究价格的意义在于预测价格，本文选取径向基（RBF）神经网络和反向传播（BP）神经网络预测黄金价格。用 BP 神经网络和 RBF 神经网络模型进行价格的模拟和预测，并用测试集的预测结果与真是的价格进行对比，得出结论。

第七章、总结与展望。本章属于总结性的章节，通过对全文的总结概括得出相应的结论，并且指出个人认为的可以改进的地方。

## 第二章、黄金市场介绍

### 2.1 黄金市场发展历史

黄金市场的发展历史主要是分为四个阶段<sup>[25]</sup>。第一阶段是十九世纪之前，即金本位制之前。由于开采冶炼技术不成熟，黄金相对稀缺，主要是封建贵族阶层所有。白银等其他贵金属充当了货币的角色。第二阶段是十九世纪初到二战结束之前的金本位制度时期。金本位制是由牛顿提出的，它是以黄金为本位币的货币制度。金本位制下每一单位的货币价值等同于若干质量的黄金。黄金作为商品交换活动中的一般等价物，行使货币职能。第三阶段是十九世纪四十年代到七十年代的布雷顿森林体系时期。此时期各国货币和美元挂钩，美元与黄金挂钩。美元替代了黄金在世界货币体系中的作用，确立了美元在世界货币体系中的重要地位。各国之间禁止黄金的自由买卖，黄金流通受到诸多限制，黄金的市场机制难以发挥作用。后来随着各国经济形势的不断演变，美国在世界经济体系中的霸主地位有所下降，美元贬值，各国开始大量兑换黄金。美国政府在 1971 年 8 月 15 日宣布停止履行外国政府和央行以美元兑换黄金的义务，布雷顿森林体系开始慢慢瓦解，黄金市场过渡到第四个阶段：信用本位时期。这一时期，黄金市场重新开放，黄金与货币逐渐脱离，黄金的市场机制开始在价格中发挥决定作用。我们平时看到的学者对于黄金价格的研究，数据也都是取自信用本位时期。

### 2.2 国际主要黄金市场

#### 2.2.1 伦敦市场

伦敦黄金交易市场具有 300 多年的发展历史。1804 年，伦敦取代荷兰阿姆斯特丹成为世界黄金交易的中心；1919 年伦敦金市正式成立，经营黄金现货交易；1982 年 4 月，伦敦期货黄金市场开业，开始经营黄金期货产品。伦敦黄金市场对黄金定盘价是独一无二的，与其他黄金市场不同，它为市场的交易者买入或卖出黄金只提供单一的报价。伦敦黄金交易市场提供的标准价格，被广泛地应用于生产商、消费者和中央银行作为中间价。

#### 2.2.2 苏黎世市场

苏黎世黄金市场是地位仅次于伦敦的世界第二大黄金市场，兴起于二战之后，没有正式组织结构，由瑞士三大银行：瑞士银行、瑞士信贷银行和瑞士联合银行负责清算结账、三大银行不仅可为客户代行交易，而且黄金交易也是这三家银行本身的主要业务。苏黎世黄金总库建立在瑞士三大银行非正式协商的基础上，不受政府管辖，作为交易商的联合体与清算系统混合体在市场上起中介作用。苏黎世黄金市场无金价定盘制度，在每个交易日

任一特定时间，根据供需状况议定当日交易金价，这一价格为苏黎世黄金官价。全日金价在此基础上的波动不受涨跌停板限制。

### 2.2.3 美国市场

美国市场是在 20 世纪 70 年代布雷顿森林体系逐渐瓦解后形成的黄金交易市场，有纽约商品交易所和芝加哥商品交易所形成。是目前世界上最大的黄金期货交易中心，对黄金现货价格有重大的影响。

### 2.2.4 中国市场

中国市场只要是上海黄金交易所，上海黄金交易所成立于 2002 年 10 月 30 日。上海黄金交易所的交易方式采用会员自主报价，以“价格优先，时间优先”的原则集中撮合成交的方式为主，辅以询价等其它交易方式。会员可选择现场方式或远端方式进行黄金交易交易所建立初期，主要开展现货交易，随着市场功能作用的发挥以及市场运行机制的完善，经国务院有关主管部门批准，开展黄金期货交易业务。本文选取的黄金价格数据就是取自上海黄金交易所的黄金现货价格。

## 2.3 黄金价格的影响因素

结合综述部分所述以及其他国内外相关研究，影响黄金价格的因素主要可以分为三类：供给因素、需求因素和其他市场因素。

### 2.3.1 供给因素

供给因素主要包括地球上的黄金存量、黄金年生产量、开采黄金的成本、主要黄金生产国的生产状况等。黄金存量越大，黄金价格越低，但黄金作为一种重要的自然资源，存量增长主要是随着生产不断增加，大约按照每年 2% 的增速增长，属于相对稳定因素。年生产量在经济相对稳定时期基本变化不大，不会对黄金的供给造成很大冲击。供给因素中，黄金的开采成本对黄金价格有着重大的影响，本世纪以来，黄金开采成本大幅度上升带动了黄金价格的上涨<sup>[26]</sup>。主要黄金生产国的生产状况则属于政治经济因素，当政治经济稳定，生产状况好的时候，会拉低国际黄金的成本，增加供给，从而拉低黄金价格。

### 2.3.2 需求因素

需求因素主要包括商业对黄金的需求量、投资需求和投机需求。在商业需求方面，黄金是一种重要的金属材料，不仅应用在首饰行业，还应用在很多其他的领域。例如微电子领域、医学领域以及建筑装饰领域等。尽管科技的进步使得黄金替代品不断出现，但黄金以其特殊的金属性质使其需求量仍呈上升趋势。商业需求的加大，推动了黄金价格的提高。在投资需求方面，黄金是一种很好的保值增值预防通货膨胀的投资品，各国央行通常

都会采用持有黄金储备的方式来防范通货膨胀。对于投资者，黄金相对于其他的一些货币资产更为稳定，可以作为预防风险的投资品。在投机需求方面，对冲基金和证券公司都在国际黄金市场上进行大规模的交易交易，形成很大的投机性需求。这些投资机构在国际黄金价格波动中扮演了重要的角色，当他们频繁买进黄金的时候，会拉高黄金价格；当抛售时，又会压低黄金价格。总体来说对黄金的需求越大，黄金价格也就越高。

### 2.3.3 其他市场因素

其他市场因素主要有美元汇率、各国货币政策、通货膨胀、股市行情、石油价格等。以前的学者对于黄金价格影响因素的研究也主要在以上这几个方面。

**美元汇率影响。**美元汇率也是影响金价波动的重要因素之一。一般在黄金市场上有美元涨则金价跌，美元降则金价扬的规律。美元坚挺一般代表美国国内经济形势良好，美国国内股票和债券将得到投资人竞相追捧，黄金作为价值贮藏手段的功能受到削弱；而美元汇率下降则往往与通货膨胀、股市低迷等有关，黄金的保值功能又再次体现。近十年最直观的例子表现为：经济危机之后，美国经济低迷，美元走低，国际黄金价格上涨，曾一度到达 1920.8 美元/盎司的历史最高点；随后伴随着美国经济复苏，美元汇率上涨，国际黄金价格从 2012 年开始连续下跌。

**各国的货币政策。**当某国采取宽松的货币政策时，由于利率下降，该国的货币供给增加，加大了通货膨胀的可能，会造成黄金价格的上升。如 60 年代美国的低利率政策促使国内资金外流，大量美元流入欧洲和日本，各国由于持有的美元净头寸增加，出现对美元贬值的担心，于是开始在国际市场上抛售美元，抢购黄金，并最终导致了布雷顿森林体系的瓦解。

**通货膨胀。**通货膨胀对黄金价格的影响要做长期和短期来分析，并要结合通货膨胀在短期内的程度而定。从长期来看，每年的通胀率若是在正常范围内变化，那么其对金价的波动影响并不大；只有在短期内，物价大幅上升，货币的单位购买力下降，金价才会明显上升。由于通货膨胀率的指标通常都是按时段公布的，例如我国 CPI 通常都是按月公布，而黄金市场的黄金价格是实时，因此短期的通过膨胀情况通常可以从汇率等相关指标中反映出来。从综述上所述研究来看，众多研究者普遍认为黄金是对抗通货膨胀的有效工具。

**股市行情。**一般来说股市下挫，金价上升。这主要体现了投资者对经济发展前景的预期，如果大家普遍对经济前景看好，则资金大量流向股市，股市投资热烈，金价下降。反之亦然。

**石油价格。**作为国际上最重要的大宗商品，石油价格的上涨通常意味着通货膨胀

的到来，黄金本身作为通货膨胀之下的保值品，因此黄金价格通常会伴随着石油价格的上涨而上涨。

除了上述影响金价的因素外，国际金融组织的干预活动，该国和地区的中央金融机构的政策法规，也将对世界黄金价格的变动产生重大的影响。

## **2.4 本章小结**

本章主要介绍了黄金市场的相关知识和影响黄金价格的相关因素。黄金市场的相关知识包括国际黄金市场的历史发展状况，国际上主要的黄金市场的简单介绍。黄金影响因素则主要从市场供给因素、需求因素和其他市场因素进行介绍，其中影响黄金价格的其他市场因素是模型选取输入变量的重要参考因素。

### 第三章、BP 和 RBF 神经网络模型简介

本章将简要介绍神经网络的发展过程、特点和功能、连接形式以及 BP 神经网络和 RBF 神经网络的相关知识。更加详细的内容请参看《前馈神经网络分析与设计》<sup>[27]</sup>、《混沌系统的模糊神经网络控制理论与方法》<sup>[28]</sup>、《前馈神经网络及其应用》<sup>[29]</sup>、《人工神经网络基础》<sup>[30]</sup>等相关书籍。

#### 3.1 神经网络模型发展过程

神经网络研究兴起于十九世纪末,最早开始美国心理学家 W. James 发表了讨论人脑结构和功能的文章,该文章对神经的学习功能进行了详细介绍,开创性的对神经网络进行了讨论<sup>[31]</sup>,由此神经网络研究正式开始。早期的神经网络研究主要集中在医学和心理学领域,在 1943 年美国生理学家 W. S. McCulloch 和数学家 W. Pitts 开始开创性的用数学工具研究神经网络,并把客观事件在神经中的信息处理用数学形式表达出来,提出了 M-P 模型<sup>[32]</sup>。尽管现在看来该模型过于简单,在权值处理上有缺陷,观点也不是很正确,但是 M-P 模型开创了神经网络研究的新方式,为以后的工作提供了依据。自二十世纪五十年代,随着计算机的发展,神经网络理论具备了初步的模拟实验条件,神经网络理论得到了较大的发展,但仍然局限在线性领域。对于非线性问题的束手无策以及人工智能的发展,使得神经网络研究在二十世纪六十年代进入萧条期。进入八十年代之后,神经网络的研究开始复兴并得到进一步的发展。1982 年,美国的 J. J. Hopfield 教授将神经网络理论与动力学系统稳定性分析结合,引入了“计算能量函数”,提出了信息存储在网络神经元的连接上的观点。他的理论为以后的神经网络研究起到了重大的作用,被称为 Hopfield 模型<sup>[33]</sup>。进入二十一世纪以来,神经网络理论迅速发展,在各个领域神经网络模型的应用都取得不错的成绩,涌现出来很多优秀的学者和好的理论。2000 年新加坡 R. Setiono 教授的快速规则抽取算法<sup>[34]</sup>;瑞士的 J. Schmidhuber 教授运用递归神经网络和深度前馈神经网络等相关网络模型做出了很多突破性研究并得到多项国际大奖。神经网络研究已进入稳定并且迅速发展的时期,运用神经网络进行研究的领域不断拓展,在智能控制、图像处理、模式识别、非线性预测等诸多领域取得突破。在经济学领域,神经网络模型已经被用来预测诸如股票、外汇等金融产品的价格。目前的神经网络模型已经有三十多种,比较具有代表性的有:支持向量机(SVM)网络、径向基函数(RBF)网络、极限学习机(ELM)、递归神经网络(RNN)、回声状态网络(ESN)等。在实际的生活,神经网络运用于计算机算法集成的超级计算程序已经在某些被认为是高难度的领域超越了人脑,诸如国际象棋。

## 3.2 神经网络的特点和功能

神经网络是仿人脑设计的网络模型，因此有类似人脑的特点和功能。这也是神经网络与其他模型的主要区别。

主要特点包括：

(1) 并行处理，可以同时执行多于一种处理的方式。神经网络中模拟人脑功能，各个神经元处理信息时相互独立，可以分别接收信息，产生结果。这种并行处理能力使得神经网络可以快速、实时处理信息。

(2) 分布式存储。传统串行体系中各个信息存储点相互影响，一个贮存单元被破坏，将导致整个系统受到影响。神经网络采用并行处理的方式，结构上采用分步式进行存储，信息存储在神经元的连接上，单个连接的权值和神经元发生变化，对最后的结果影响都不大，这也造就了神经网络极强的鲁棒性。

(3) 鲁棒性，指在收到干扰的情况下维持稳定的能力。神经网络的信息主要体现在神经元之间的连接权值上，神经网络存在大量的连接，当个别的权值改变的，不会对系统原有信息造成很大的影响，因此具有很强的鲁棒性。

(4) 可塑性和学习能力。人工神经网络可以通过训练和反馈不断的调整权值，使输入时得到期望的输出值。

神经网络功能主要包括：

(1) 非线性映射。神经网络是一种解决高度非线性问题的有效模型，可以高效的建立输入和输出层之间的非线性关系模型，对大多数非线性数据有较好的模拟效果。

(2) 联想存储功能。由于神经网络本身具有分布式存储和并行处理的特点，神经网络通过各个信息单元协同处理信息，因此神经网络模型具有联想存储功能。

(3) 最优化问题计算功能。最优化是指在特有的约束条件下找到待解决目标的最优解。神经网络模型是非线性的模型，网络处理问题过程中向最优化值靠近的过程就是一个最优化问题的求解过程。

(4) 知识获取和处理能力。知识的获取和处理能力是人脑的重要功能之一，神经网络模型是仿人脑结构的数学模型，在知识的获取和处理能力上面有着比较强的能力。神经网络模型通过对输入和输出信息进行抽取，并且将抽取处理后的结果储存在网络中，也正是因为这种能力，使得神经网络模型可以在没有先验知识的情况下从输入的数据中提取有用的特征信息，发现规律进行预测。

### 3.3 神经网络常见的连接形式

随着神经网络模型不断发展，截至目前，已经有几十种神经网络模型被研究出来，其中以反馈网络模型和前馈网络模型比较常见和典型，这里简要回顾一下这两种连接形式神经网络模型的网络结构。

#### 3.3.1 反馈网络模型

反馈神经网络的网络结构如下图 3-1 所示，该网络在输入层和输出层之间存在反馈调节，因此被成为反馈神经网络，不同于前馈神经网络的是，反馈神经网络的每一个输入及节点都可能接受外部的输入和来自内部输出神经元的反馈

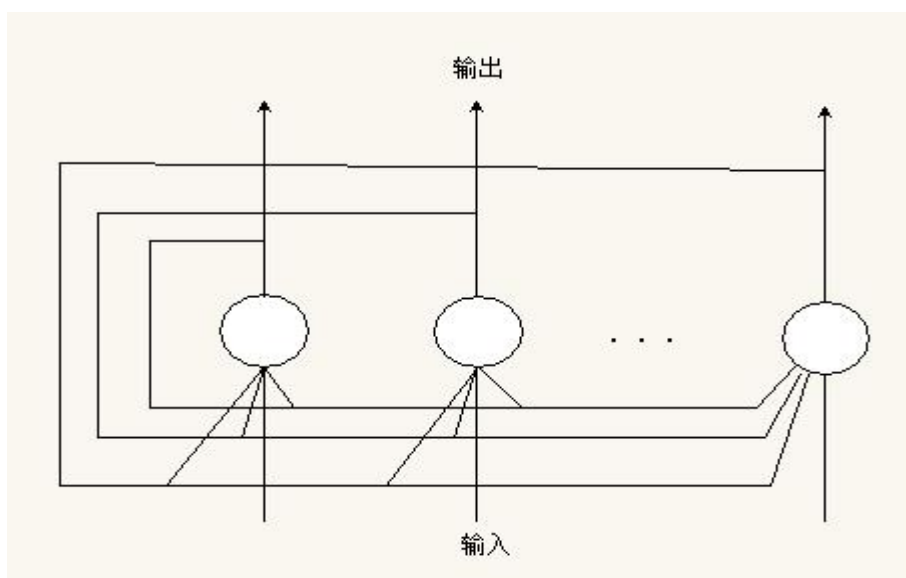


图 3-1 反馈型神经网络结构图

#### 3.3.2 前馈神经网络模型

神经元分层排列，有输入层，隐含层（可以是多个）和输出层三部分构成，信息是由上一层传递到下一层，同一层之间不进行信息传递，前馈型神经网络结构图为如下图 3-2 所示的有向无环图。

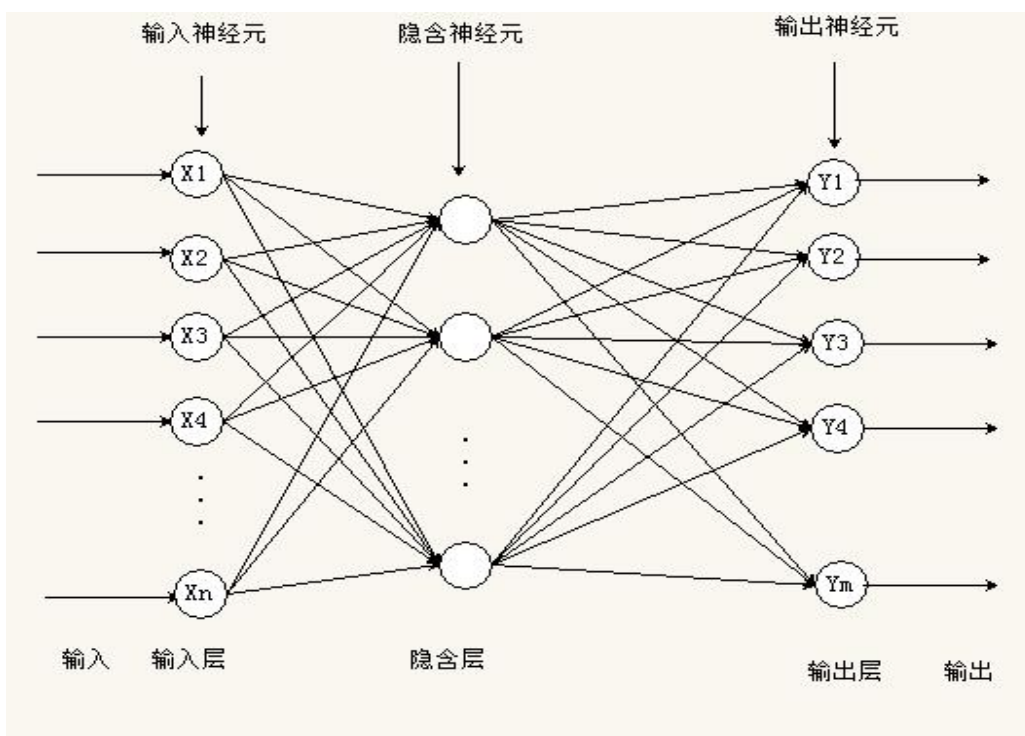


图 3-2 前馈型神经网络结构图

目前的应用比较广范的前馈型神经网络有反向传播（BP）神经网络和径向基函数（RBF）神经网络等。本文将用 BP 神经网络和 RBF 神经网络预测黄金价格。

### 3.4 神经网络的学习方法

学习是神经网络的一个重要特性，神经网络可以通过学习算法，实现自我学习能力。目前神经网络的学习算法已经有很多种，按照有无导师可以分类成有导师指导的神经网络学习、无导师指导的神经网络学习和增强式神经网络学习。

在有导师指导的神经网络学习算法中，网络根据输入计算的输出会和期望的输出进行比较，然后根据差异调整神经网络的权值，最终使差异最小化，具体的结构如下图 3-3。。

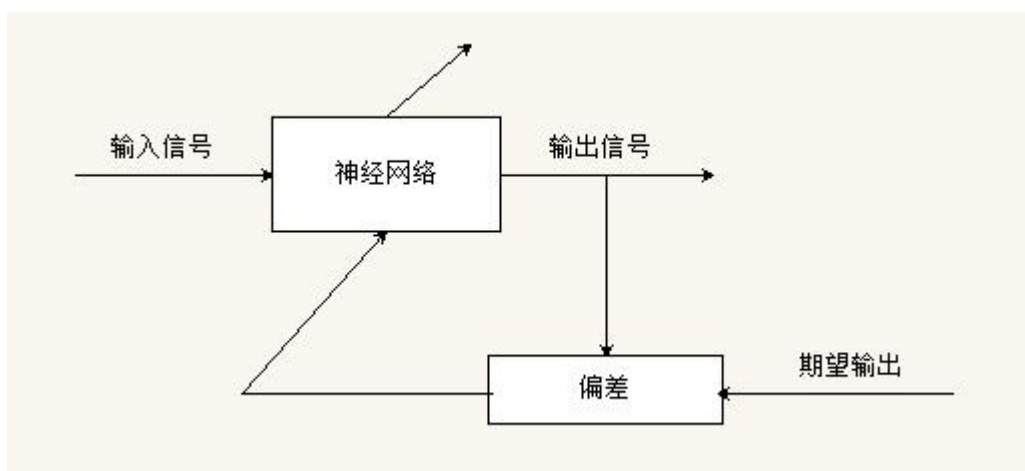


图 3-3 有导师指导的学习算法结构图

在无导师的神经网络学习算法中，网络是按照事先设定的规则自动调整权值，并没有将输出和期望输出进行对比反馈，具体的学习方式如下图 3-4。

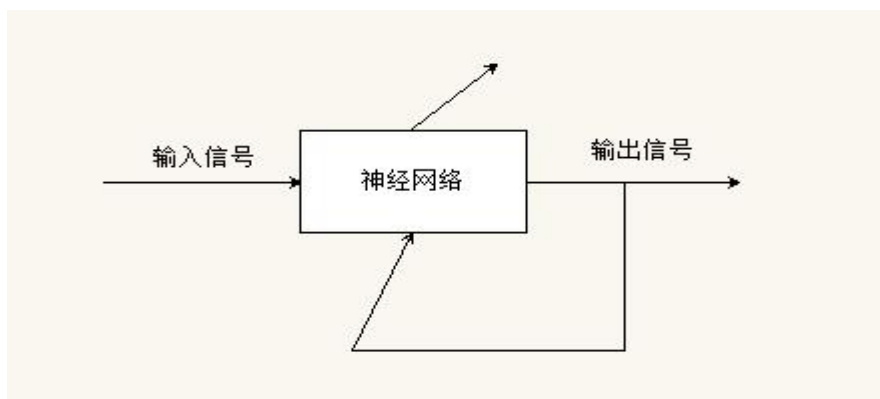


图 3-4 无导师指导的学习算法结构图

增强式神经网络学习算法就是一种介于有导师神经网络学习算法和无导师神经网络学习算法之间的一种学习方式。

本文选用的 BP 神经网络模型和 RBF 神经网络模型在实际预测的时候，都是选用的有导师学习的算法进行模拟预测的。

### 3.5 BP 神经网络

1974 年 Werbos 提出有关 BP 神经网络学习算法以及几种变形形式<sup>[35]</sup>，1986 年 Rumelhart 和 McClelland 正式提出 BP 神经网络<sup>[36]</sup>，是目前应用最广泛的神经网络模型之一。BP 神经网络由多个神经元形成，具有非常强的非线性处理能力，是以一种前馈型、有导师指导的神经网络学习模型，主要由输入信号的正向传播和误差信号的反向传播构成。标准的 BP 神经网络结构如图 3-5，包括输入层，隐含层和输出层，正向传播通过输入层输入，并经过非线性的映射通过隐层输出；误差信号的反向传播是通过将训练值和模型输出结果进行比较，将误差逐层向前反馈，不断调整各层连接的阈值和权值。

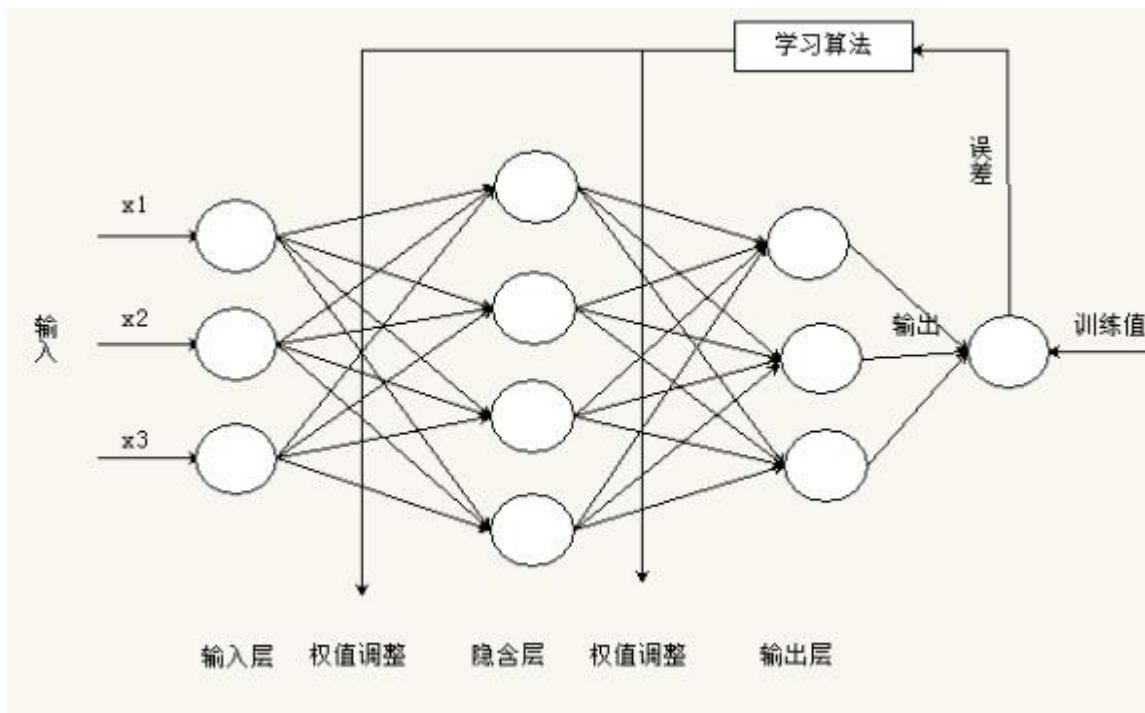


图 3-5 BP 神经网络结构图

设输入层包含  $N$  个神经元节点，输入向量为  $u = (u_1, u_2, \dots, u_N)^T \in R^N$ ，隐含层包括  $L$  个神经元节点， $\varphi(\bullet)$  代表隐含层神经元的激活函数， $\theta_j, (j=1, 2, \dots, L)$  代表隐层神经元的阈值，输出层包含  $M$  个神经元节点，输出向量为， $y = (y_1, y_2, \dots, y_M)^T \in R^M$ ， $\psi(\bullet)$  表示输出层神经元激活函数， $\theta_k, (k=1, 2, \dots, M)$  表示输出层神经元阈值， $w_{ij}$  表示输入层神经元到隐含层神经元之间的连接权值， $w_{jk}$  表示隐含层神经元到输出层神经元之间的连接权值。

### 3.5.1 信号正向传播过程

正向传播过程主要是指输入信息从输入层经过隐含层最后到达输出层的过程。设隐含层第  $j$  个神经元节点的输入信号为  $net_j$ ，则

$$net_j = \sum_{i=1}^N w_{ij} u_i + \theta_j$$

神经元通过激活函数非线性传导，设隐含层第  $j$  个神经元节点的输出信号为  $o_j$ ，则

$$o_j = \varphi(net_j) = \varphi\left(\sum_{i=1}^N w_{ij} u_i + \theta_j\right)$$

各个层之间通过权值线性连接，设输出层第  $k$  个神经元节点的输入信号为  $net_k$ ，则

$$net_k = \sum_{j=1}^L (w_{jk}o_j + \theta_k) = \sum_{j=1}^L (w_{jk}\varphi(\sum_{i=1}^N w_{ij}u_i + \theta_j) + \theta_k)$$

输出层到第  $k$  个神经元节点的输出信号为  $y_k$ ，则

$$y_k = \psi(net_k) = \psi(\sum_{j=1}^L (w_{jk}\varphi(\sum_{i=1}^N w_{ij}u_i + \theta_j) + \theta_k))$$

### 3.5.2 误差的反向传播过程

误差的反向传播过程是一个有导师指导的学习过程，根据信号的正向传播输出信号与实际值（期望值）进行比较，计算输出层输出值和实际值之间的误差。将误差从输出层开始逐层向前传递，通过梯度下降算法修改各层之间的连接权值和阈值，通过调整，使得最后输出值和实际值之间的误差最小。具体的算法如下：

设输入值包括  $P$  个样本，每个训练样本的值为  $x^{(p)}$ ，输出层第  $k$  个神经元节点的实际值（期望输出）和实际输出分别为  $d_k^{(p)}$  和  $y_k^{(p)}$ ，对每一个训练样本  $x^{(p)}$  ( $p=1,2,\dots,P$ )，误差的二次型函数为  $E^{(p)}$ ，输出层和隐含层的误差信号分别为  $\delta_k$  和  $\delta_j$ ，则

$$E^{(p)} = \frac{1}{2} \sum_{k=1}^L (d_k^{(p)} - y_k^{(p)})^2$$

包含  $P$  个样本值的训练样本的误差综合为  $E$ ，则

$$E = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^M (d_k^{(p)} - y_k^{(p)})^2$$

根据梯度下降算法向前一层一层调整 BP 神经网络的连接权值和阈值，设隐含层传递到输出层的连接权值修正值为  $\Delta w_{jk}$ ，输出层阈值的修正值为  $\Delta \theta_k$ ，输入层传递到隐含层的连接权值的修正值为  $\Delta w_{ij}$ ，隐含层阈值的修正值为  $\Delta \theta_j$ ，学习速率为  $\eta$ 。

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}}; \Delta \theta_k = -\eta \frac{\partial E}{\partial \theta_k}; \Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}}; \Delta \theta_j = -\eta \frac{\partial E}{\partial \theta_j}$$

隐含层传递到输出层连接权值的调整公式为

$$\Delta \theta_k = -\eta \frac{\partial E}{\partial \theta_k} = -\eta \frac{\partial E}{\partial net_k} \frac{\partial net_k}{\partial \theta_k} = -\eta \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial net_k} \frac{\partial net_k}{\partial \theta_k}$$

输出层阈值的调整公式为

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \frac{\partial E}{\partial net_k} \frac{\partial net_k}{\partial w_{jk}} = -\eta \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial net_k} \frac{\partial net_k}{\partial w_{jk}}$$

输入层传递到隐含层的连接权值的调整公式为

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = -\eta \frac{\partial E}{\partial net_j} \frac{\partial net_j}{\partial w_{ij}} = -\eta \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ij}}$$

隐含层的阈值调整公式为

$$\Delta \theta_j = -\eta \frac{\partial E}{\partial \theta_j} = -\eta \frac{\partial E}{\partial net_j} \frac{\partial net_j}{\partial \theta_j} = -\eta \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial net_j} \frac{\partial net_j}{\partial \theta_j}$$

因为 
$$\frac{\partial E}{\partial y_k} = -\sum_{p=1}^P \sum_{k=1}^M (d_k^{(p)} - y_k^{(p)})^2$$

$$\frac{\partial net_k}{\partial w_{jk}} = o_j; \frac{\partial net_k}{\partial \theta_k} = 1; \frac{\partial net_j}{\partial w_{ij}} = u_i; \frac{\partial net_j}{\partial \theta_j} = 1$$

$$\frac{\partial E}{\partial o_j} = -\sum_{p=1}^P \sum_{k=1}^M w_{jk} (d_k^{(p)} - y_k^{(p)}) \psi'(net_k)$$

$$\frac{\partial y_k}{\partial net_k} = \psi'(net_k); \frac{\partial o_j}{\partial net_j} = \varphi'(net_j)$$

通过以上过程可以得到如下公式

$$\Delta w_{jk} = \eta \sum_{p=1}^P \sum_{k=1}^M (d_k^{(p)} - y_k^{(p)}) \psi'(net_k) o_j = \eta \delta_k o_j$$

$$\Delta \theta_k = \eta \sum_{p=1}^P \sum_{k=1}^M (d_k^{(p)} - y_k^{(p)}) \psi'(net_k) = \eta \delta_k$$

$$\Delta w_{ij} = \eta \sum_{p=1}^P \sum_{k=1}^M w_{jk} (d_k^{(p)} - y_k^{(p)}) \psi'(net_k) \varphi'(net_j) u_i = \eta \delta_j u_i$$

$$\Delta \theta_j = \eta \sum_{p=1}^P \sum_{k=1}^M w_{jk} (d_k^{(p)} - y_k^{(p)}) \psi'(net_k) \varphi'(net_j) = \eta \delta_j$$

在实际应用神经网络模型时，激活函数通常有以下几种：

**Sigmoid 函数：**

$$f(x) = \frac{1}{1 + e^{-x}}$$

**双曲正切函数：**

$$f(x) = \tanh x = \frac{\sinh x}{\cosh x}, \text{其中 } \sinh x = \frac{e^x - e^{-x}}{2}, \cosh x = \frac{e^x + e^{-x}}{2}$$

反正切函数:

$$f(x) = \arctan x$$

自适应多项式激活函数

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

### 3.6 RBF 神经网络

径向基函数 (RBF) 神经网络是 1988 年由 D. Broomhead 和 D.Lowe 提出来的一种前馈型神经网络<sup>[37]</sup>。RBF 一共由输入层, 隐含层, 输出层三部分组成, 和 BP 神经网络不同的是 RBF 神经网络的隐含层节点激活函数是径向基函数, 只调整输出权值, 因此学习速度较快, 并且可以避免局部的最小值。

#### 3.6.1 RBF 神经网络的网络结构

RBF 神经网络是三层式的前馈神经网络, 第一层是输入层, 输入层节点的个数等于输入数据的维数, 是直接和外界输入数据相连接的层, 设输入向量为  $x, x = (x_1, x_2, \cdots, x_N)$ ,  $N$  为输入神经元节点个数; 第二层是隐含层, 隐含层的节点个数需要通过具体的模型和数据特点进行判断, 隐含层通过径向基函数将数据从输入层通过径向基函数的非线性变化形成新的数据; 第三层是输出层, 一般采用线性激活函数, 通过将隐含层通过径向基函数生成的输出信号线性求和得到输出值, 设连接隐含层和输出层之间的权值向量为  $w, w = (w_1, w_2, \cdots, w_M)^T$ 。具体的网络结构如下图 3-6。

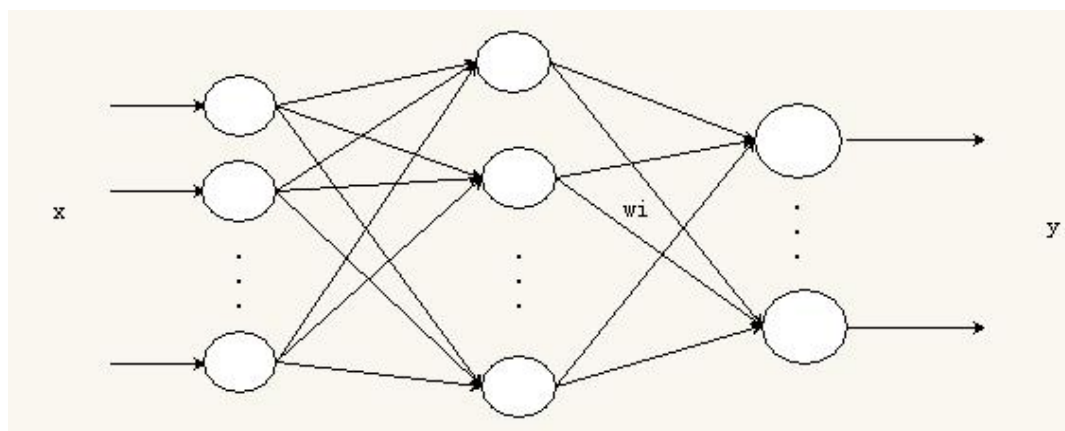


图 3-6 RBF 神经网络结构图

### 3.6.2 径向基函数

径向基函数是一种取值只依赖于距离的函数。一般情况下径向基函数记为  $\Phi(x, c) = \varphi(\|x - c\|)$ ，其中  $\|\bullet\|$  为欧几里德范数。径向基函数满足：如果  $\|x_1\| = \|x_2\|$ ，则  $\Phi(x_1) = \Phi(x_2)$ 。径向基函数是局部分布的对中心点径向对称衰减的非负非线性函数，常见的径向基函数一般是一下三种：

Kriging 方法的高斯函数：  $\varphi(r) = e^{-\frac{r^2}{\sigma^2}}$

反演的 Sigmoid 函数：  $\varphi(r) = \frac{1}{1 + e^{\frac{r^2}{\sigma^2}}}$

逆多二次函数：  $\varphi(r) = \frac{1}{(r^2 + \sigma^2)^{\frac{1}{2}}}$

其中  $\sigma$  是求解的径向基函数宽度

### 3.6.3 K 均值聚类算法

在 RBF 神经网络模型中，由于径向基函数是距离的函数，因此在构造模型的过程中，选取合适的初始数据中心非常重要，本文介绍的是 Moody 和 Darken 在 1989 年提出来的 K 均值聚类算法来求解数据中心<sup>[38]</sup>。K 均值聚类的基本方法是首先选定聚类中心个数为 M，即设计的神经网络模型的隐含层神经元个数为 M。选取 M 个聚类中心后，设定最后的聚类中心为  $c_1, c_2, \dots, c_M$ ；算法的最大迭代次数为 T；第 t 次迭代之后形成的聚类中心分别为  $c_j(t), j = 1, 2, \dots, M$ ，相应的聚类点形成的聚类域用  $U_j(t)$  表示。则 K 均值聚类的具体算法步骤如下：

首先初始化聚类中心。开始随机选取 M 个不同的点作为初始径向基函数的聚类中心，表示为  $c_j(0), j = 1, 2, \dots, M$ 。然后计算输入样本  $x^{(p)}, p = 1, 2, \dots, M$  与第一步随机生成的聚类中心之间的距离，这里的距离采用的是欧式距离，即  $d = \|x^{(p)} - c_j(t)\|$ 。根据距离邻近原则，将输入值归类到与之距离最近的聚类中心的聚类域中，将所有的数据以此分类，形成相应的聚类域，即当输入样本  $x^{(p)}$  满足  $j^* = \arg \min_j \|x^{(p)} - c_j(t)\|$  时，将样本点  $x^{(p)}$  划分到聚类域  $U_{j^*}(t)$  当中。再通过计算聚类域中的样本平均值，重新计算新的聚类中心，

$c_j(t+1) = \frac{1}{M_j} \sum_{x^{(p)} \in U_j(t)} x^{(p)}$ ，其中  $M_j$  为聚敛域  $U_j(t)$  中样本点的个数。将 t 值替换成 t+1，重复

进行上步的计算，直到聚类中心  $c_j$  得变动小于预设的变动阈值或者是达到了最大的迭代次数停止。具体的流程图如下图 3-7。

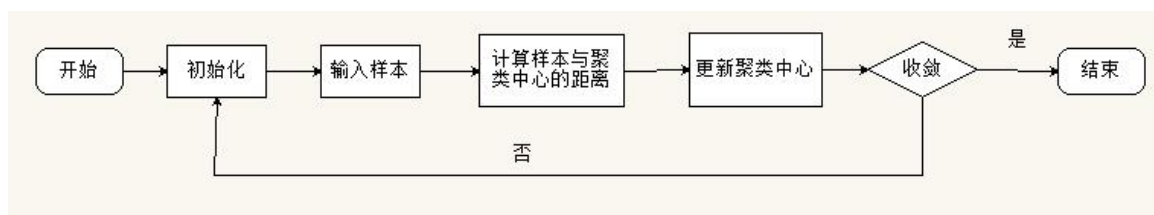


图 3-7 K 均值聚类算法流程图

### 3.6.4 正则化理论

正则化理论是 Tikhonov 在 1963 年提出来的<sup>[39]</sup>，用来解决不适定问题的一种方法。基本原理为：加入一个有解的先验知识约束来控制映射的光滑性。正则化的网络具有最佳逼近特性、最佳解以及可以逼近任意精度紧集上的任意多远连续函数的特征。

设逼近函数为  $F(X)$ ；本文中最终的输出变量为价格，是一维变量，因此假设函数的输出也是一维的，用  $y$  表示；用函数进行逼近的一组数据为：

输入数据：  $X^{(p)}, p=1,2,\dots,P$ ；期望输出：  $d^{(p)}, p=1,2,\dots,P$

传统的寻找逼近函数的方法都是通过使输出数据和期望输出的总离差平方和最小化来实现的，即使  $E_s(F) = \frac{1}{2} \sum_{p=1}^P (d^{(p)} - y^{(p)})^2 = \frac{1}{2} \sum_{p=1}^P [d^{(p)} - F(X^{(p)})]^2$  最小化。正则化方法则

是在该总离差平方和的基础上加上一个正则项，其中正则项是一个控制逼近函数光滑程度的项，正则化项为  $E_c(F) = \frac{1}{2} \|DF\|^2$ ，其中  $D$  为线性微分算子，表示  $F(X)$  的先验知识。此时

在正则化条件下求最小化变成了有约束条件的最小化，表示为

$E(F) = E_s(F) + \lambda E_c(F) = \frac{1}{2} \sum_{p=1}^P [d^{(p)} - F(X^{(p)})]^2 + \frac{1}{2} \lambda \|DF\|^2$  从公式可以看出，第一项就是通过

样本求出来的总离差平方和，由样本数据决定；第二项是有先验知识决定的。用  $F_\lambda(X)$  表示函数的最小解函数，当  $\lambda \rightarrow 0$  表明约束条件，即先验知识对最小化结果没有约束作用，最小化的解完全取决于样本数值；当  $\lambda \rightarrow \infty$  表明最小化的解函数完全可以由先验知识决定，样本不包含有用的信息。当  $\lambda$  在正无穷和零之间取值的时候，表明样本信息和先验知识对最小化的解函数都有作用。

对于正则化问题，一般情况下解都为： $F(X) = \sum_{p=1}^P w_p G(X, C)$ ，其中为 $G(X, C)$ 格林函数； $w_p$ 为权系数。格林函数 $G(X, C)$ 与 $D$ 有关，若 $D$ 满足平移不变性和旋转不变性，则格林函数表现为距离函数，即 $G(X, C) = G(\|X - C\|)$ 。多元 Gauss 就是这中函数的一个典型例子。

### 3.6.5 输出层权值的确定

本文考虑在正则化条件下权值的确定方法，本质上还是通过权值的调整使得误差平方和最小化，即求解函数 $E(F)$ 最小整数解 $F_\lambda(X)$ 的过程。

设逼近函数 $F(X) = \sum_{i=1}^M w_i G(\|x - c_i\|)$ ，这里 $c_i$ 表示第 $i$ 个聚类中心，一共 $M$ 个聚类中心

$G$ 为 Gauss 函数， $\xi(F^*) = \sum_{p=1}^P [d^{(p)} - F(X^{(p)})]^2 + \lambda \|DF^*\|^2$ ，求解函数的过程就是求解权值集

的过程。则

$$\begin{aligned} \xi(F^*) &= \sum_{p=1}^P \left[ d^{(p)} - \sum_{i=1}^M w_i G(\|x^{(p)} - c_i\|) \right]^2 + \lambda \|DF^*\|^2 = \sum_{p=1}^P (d^{(p)} - GW)^2 + \lambda (DF^*, DF^*)_H^2 \\ &= \|d - GW\|^2 + \lambda w^T G_0 w \end{aligned}$$

其中 $d = (d^{(1)}, d^{(2)}, \dots, d^{(P)})^T$ ， $w = (w_1, w_2, \dots, w_M)$

$$G = \begin{bmatrix} G(x^{(1)}, c_1) & G(x^{(1)}, c_2) & \cdots & G(x^{(1)}, c_M) \\ G(x^{(2)}, c_1) & G(x^{(2)}, c_2) & \cdots & G(x^{(2)}, c_M) \\ \vdots & \vdots & \ddots & \vdots \\ G(x^{(P)}, c_1) & G(x^{(P)}, c_2) & \cdots & G(x^{(P)}, c_M) \end{bmatrix}$$

使 $\xi(F^*)$ 最小化时满足 $(G^T G + \lambda G_0)w = G^T d$ ，可以求解出正则化你条件下的权值 $w$ ，

$$w = (G^T G + \lambda G_0)^{-1} G^T d$$

RBF 神经网络也是一种前馈型的神经网络模型，在确定输出层权值的时候使用的具体学习算法和 BP 神经网络模型的学习算法类似。通过迭代，在正则化条件下最后形成的解为

$$w_{k+1} = w_k - \eta \frac{\partial E(F)}{\partial w_k} = w_k - \eta [(G^T G + \lambda G_0)w_k - G^T d]$$

### 3.7 本章小结

本章主要介绍本文用来预测的 BP 神经网络和 RBF 神经网络模型。开始简单介绍神经网络的发展历史、特点和功能、连接形式、学习方法相关的基本概念。然后重点介绍了 BP 神经网络和 RBF 神经网络，包括 BP 神经网络的信号正向传播和误差反向传导，RBF 神经网络确定中心的 K 均值算法和正则化算法。

## 第四章、时间序列的非线性判定

基于神经网络对非线性数据具有很好的预测，而如果是到的时间序列是线性的，则可以用其他很多线性模型就行预测，因此本文首先对黄金价格序列的线性特征进行研究。而混沌又是非线性系统中一种特有的形式，在自然研究领域包括气象学、动力学、生物学等和包括经济学在内的人文领域都广泛存在。本章将判断黄金价格的时间序列是否是混沌的，如果是，就说明黄金价格的时间序列是非线性的。

### 4.1 混沌的定义、特征及判别方法

尽管混沌目前成为很多学科研究的前沿领域，但是关于混沌的学术定义，至今没有统一。在此引用 1975 年李天岩和 Yorke 提出来的混沌的定义<sup>[40]</sup> (Li-Yorke 定义)，这一定义目前已被学者广泛接受并加以修改。

Li-Yorke 定义：设连续映自映射  $f : I \rightarrow I \subset \mathbb{R}$ ,  $I$  是  $\mathbb{R}$  中的一个子区间，如果存在不可数集合  $S \subset I$  使得：

- (1)  $S$  不包含周期点
- (2) 任何  $X_1, X_2 \in S (X_1 \neq X_2)$ , 有

$$\limsup |f^t(X_1) - f^t(X_2)| > 0$$

$$\liminf |f^t(X_1) - f^t(X_2)| = 0$$

其中,  $X_1 \in S$  及  $f^t(\bullet) = f(f(\dots f(\bullet)))$  表示  $t$  重函数

- (3) 任给的任意周期点  $P \in I$ , 有

$$\limsup_{t \rightarrow \infty} |f^t(X_1) - f^t(P)| > 0$$

则称  $f$  在  $S$  上是混沌的。

从早期的 Li-Yorke 定义可以看出，子集  $I$  中的点实际上是相当分散而又集中的，同时子集也是不会趋近于周期点。但是该定义并不能保证混沌是可观测的。在实际应用过程中，人们需要的往往是可测情况，即  $S$  有一个正的测度。之后的研究人员，包括 Day、Devaney、Smale 等都给出了关于混沌的不同定义，由于相关工作者从事的研究领域不同，混沌的定义都是基于自身研究的领域形成的关于混沌现象的解释，一次至今仍然没有公认的、普适化的混沌定义。但是不论基于的领域有何区别以及定义方式有何不同，混沌现象的本质特征是相同的。总结起来主要有一下几种特征：

- (1) 混沌序列对初始值极为敏感，初始值的微小变化会对最终的结果产生重大影响。

传统理论认为两个初始值只有微小差别的系统中，系统变化的轨迹应该是相似的，但是在混沌系统中，初始值的微小差异，可以导致两种完全不同的结果。

(2) 内在随机性。即使系统不受到任何外部干扰，给定初始值，混沌系统内部也会产生随机波动。经典线性理论认为，在给定的线性系统中，通过不断地搜集系统信息，可以降低系统的不确定性，甚至当得到信息足够多的时候，可以消除系统的随机性，给定特定的初始值，得到相应的系统状态。但是在混沌理论中，由于系统存在内在随机性，因此随机性在混沌系统中是无法被消除的，也只有在混沌系统中才存在仅仅几个变量简单系统就可以产生的随机性。

(3) 遍历性。在混沌系统中，系统的运动轨迹将会混沌吸引子内的每一个状态点。

(4) 具有奇异吸引子与分形结构。奇异吸引子又称为混沌吸引子，是混沌系统中特有的一种特性，是系统内部总体稳定和局部的不稳定的一种体现。表现为：在奇异吸引子外部的运动，最后都趋近于（吸引到）奇异吸引子，在奇异吸引子内部的运动相互排斥表现的不稳定。奇异吸引子的一个比较明显的例子就是洛伦兹吸引子。分形是奇异吸引子特征的一个几何体现，混沌系统中内部具有自相似性，各种奇异吸引子都有分形结构，用分数维可以描述其特性的几何结构。

本章是基于对黄金价格时间序列的非线性特征进行判定，通过判定黄金价格的时间序列是一个混沌的时间序列，来说明黄金价格时间序列自身的非线性。当一个时间序列是混沌的，系统中就会存在混沌吸引子，混沌吸引子满足以下两个特征：一、具有自相似的分形特征；二、对初始的条件比较敏感。其中在实际运用过程中，判断混沌性的方法主要是分为定性分析和定量分析。以下将介绍定性分析和定量分析的相关算法及原理。

## 4.2 定性分析

定性分析主要包括吸引子相图法、功率谱法、代替数据 Poincare 截面法等，是根据序列在观测域内的一些相关特性进行粗略的分析，将结果与混沌序列展示出来的结果进行比较，从现象上判断序列是否是混沌的一种方法。以下介绍定性分析中功率谱法的相关理论原理。

**功率谱法。**功率谱是功率谱密度函数的简称，表示信号功率在频域的分布状况。功率谱分析法，是基于数据的统计特征来研究序列是否混沌的一种方法。

功率谱法的一般原理：一般情况下，实际得到的序列中都是按照一定时间间隔隔开的时序  $x_i, i=1,2,3,\dots,n$ 。利用频率  $f$  与对应的功率之间的关系，求自相关函数  $C_l$ ，

$t=1,2,3\cdots n$  的傅里叶变换, 求出自功率谱密度函数  $p_k$ ,  $k=1,2,3\cdots n$ ,  $p_k$  就表示第  $k$  个频率分量对  $x_t$  的贡献, 一个混沌时间序列可以看成是各个周期运动经过叠加形成的,  $p_k$  则代表了单位频率上的能量分配, 即功率谱。

具体数学算法如下: 对于一个时间序列  $x_i, i=1,2,3,\cdots,n$ , 对采集的  $n$  个值加上周期条件  $x_{n+t} = x_t$ , 可以计算出序列  $x_i$  的自相关函数  $C_t$

$$C_t = \frac{1}{n} \sum_{i=1}^n x_i x_{i+t} \quad t=1,2,3,\cdots,n$$

对自相关函数  $C_t$  进行傅里叶变换可得自功率谱密度函数  $p_k$

$$p_k = \sum_{t=1}^n C_t e^{\frac{-2\pi kt \sqrt{-1}}{n}} \quad k=1,2,3,\cdots,n$$

通过如上功率谱的计算方法可以得到一个离散时间序列功率谱的数值, 通过比较功率谱图形可以进行序列特性的分析。

根据功率谱理论, 当一个序列呈现周期或者是逆周期特性的时候, 其对应的功率谱是分立式的, 具有离散的尖峰; 当一个序列是随机过程, 那么该序列的在各个频率下对应的振幅和频率本身是不相关的, 所有的能量在各个频率内均匀分布, 功率谱密度在整个区域内也是均匀分布的, 功率谱图上表现出的特点是连续平稳的; 当一个时间序列是混沌的, 混沌系统中具有确定性的随机性, 其形成的功率谱图是连续的, 但与随机序列不同的是混沌序列生成的功率谱图并不是平稳的, 通常都会有噪音背景, 整体趋势一般是随着频率的增高而下降。

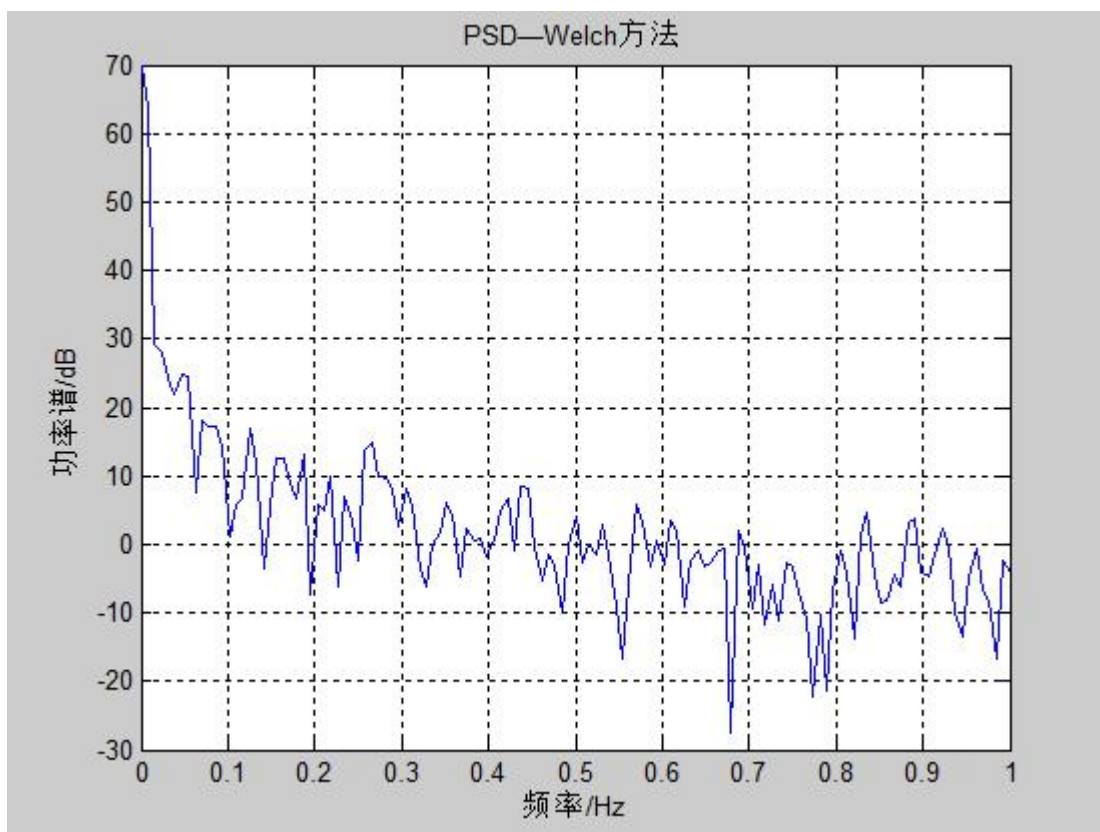


图 4-1 黄金价格功率谱图

通过功率谱法绘制的黄金价格的功率谱图如图 4-1，从图中可以看出黄金价格功率谱图是连续的、不平稳的、有噪音背景的，并功率随着频率的增加而减小。满足混沌序列频率谱图的特点，因此根据频率谱图判断黄金价格的时间序列是混沌的。因此可以判断黄金价格是非线性的。

### 4.3 定量分析

定量分析是指对客观对象的数量特征、数量关系以及数量变化的进行分析，得出结果的分析方法。在此针对混沌时间序列的定量分析就是指通过计算混沌序列的某些特征量，结合混沌序列特有特征进行判断的一种方法。如果时间序列是混沌的，那么由此序列计算出来的定量指标应当满足对应的条件，因此可以根据定量指标的值来判断时间序列是否是混沌的。常用的混沌序列判别的定量分析方法主要有关联维数法、Lyapunov 指数法、Kolmogorov 熵，本文选取 Lyapunov 指数法对时间序列进行分析。

#### Lyapunov 指数法

混沌系统具有对初始值敏感的特征，Lyapunov 指数是一个计算初始相近轨道随时间推移的分离程度的指数。只要最大 Lyapunov 指数大于 0，系统就是混沌的。

设时间序列  $x_0, x_1, \dots, x_n$  迭代方程的一般形式为  $x_n = f(x_{n-1})$ ， $\delta_i$  为第  $i$  次迭代之后产生的

偏差。因为混沌序列对初始值敏感，设定初始值为  $x_0$ ，在初始值上施加一个微小的扰动  $\delta_0$ ，

$$\text{则经过第一次迭代之后 } x_1 + \delta_1 = f(x_0 + \delta_0)$$

$$\text{经过第二次迭代后 } x_2 + \delta_2 = f(x_1 + \delta_1) = f^{(2)}(x_0 + \delta_0)$$

$$\text{以此类推，经过 } n \text{ 次迭代之后 } x_n + \delta_n = f(x_{n-1} + \delta_{n-1}) = f^{(n)}(x_0 + \delta_0)。$$

$$\text{因为 } x_n + \delta_n = f^{(n)}(x_0 + \delta_0) = f^{(n)}(x_0) + \left(\prod_{i=0}^{n-1} f'(x_i)\right)\delta_0, \quad x_n = f^{(n)}(x_0)$$

$$\text{所以由于初始值微小变化最终到时 } n \text{ 次迭代之后的差异值 } \delta_n = \left(\prod_{i=0}^{n-1} f'(x_i)\right)\delta_0$$

$$\text{即 } \left|\frac{\delta_n}{\delta_0}\right| = \left|\prod_{i=0}^{n-1} f'(x_i)\right|, \text{ 对等式进行如下恒等变换}$$

$$\left|\frac{\delta_n}{\delta_0}\right| = \left|\prod_{i=0}^{n-1} f'(x_i)\right| = e^{\ln \left|\prod_{i=0}^{n-1} f'(x_i)\right|} = e^{n \left[\frac{1}{n} \ln \left|\prod_{i=0}^{n-1} f'(x_i)\right|\right]} = e^{n \left[\frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)|\right]} = e^{n \lambda_n}$$

$$\text{其中 } \lambda_n = \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)|$$

$$\text{定义 } \lambda = \lim_{n \rightarrow \infty} \lambda_n = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)| \text{ 为 Lyapunov 指数}$$

本文中所用的黄金价格时间序列经过 MATLAB 算法编译的 Wolf 方法<sup>[41]</sup>求解出来的  $\lambda_{\max}$  为 0.0647。可以判断黄金价格时间序列是混沌的。因此可以判断黄金价格的时间序列是非线性的。

#### 4.4 本章小结

本章的主要目的介绍是介绍混沌现象的判别方法和相关的算法，通过判断混沌性，说明黄金价格序列是混沌的，进而说明黄金价格时间序列是非线性的。为了进行混沌特性判别，首先从混沌的定义开始介绍，并且介绍混沌系统中具有的相关性质和特点，然后通过混沌的系统的特点来进行混沌特性的判别。判别混沌理论的方法从定性判别和定量判别两方面展开，主要是介绍了定性判别的功率谱法和定量判别的最大 Lyapunov 指数法，并且根据收集的黄金价格数据进行定性和定量的两方面判定，得到黄金价格是混沌序列的结果。通过混沌性质的判定，表明黄金价格数据具有非线性特征。

## 第五章、时间序列的空间重构

在神经网络模型中，神经元个数的选取一直都是个很难解决的问题，传统方法上多采用经验方法，或者通过实验的方法，选择最优的神经元个数。通过实验的方法虽然可以选择到最优的神经元个数，但是在数据量很大的时候，需要进行多次实验的选取，明显是非常耗费时间的，本文拟通过相空间重构确定的最小嵌入维数作为神经元个数进行神经网络模型预测。

在实际的研究中，通常高维的序列我们只能得到一个一维信息，比如黄金价格的时间序列，显然黄金价格是受多种因素共同影响的结果，但是得到的结果仅仅是价格和时间的关系。空间重构就是通过一维的时间序列反向构造出具有相同性质的多维空间来描述一维时间序列中隐藏的规律。通过空间的重构，构造一个具有和原来系统相似性质的新系统来对原来的系统进行分析。

### 5.1 Takens 定理

荷兰数学家 Takens 提出的嵌入定理<sup>[42]</sup>证明了当嵌入的空间维数  $m$ ，满足  $m \geq 2d + 1$  时， $d$  为原系统的维数是，可以通过空间重构还原混沌吸引子的轨道，构造一个与原来混沌系统微分同胚的新系统。

在描述 Takens 定理之前先介绍等距同构和嵌入两个相关的概念

等距重构：设存在两个度量空间  $(N_1, \rho_1)$  和  $(N_2, \rho_2)$ ，如果存在一个映射  $\phi$  使得 (1)  $\phi$  是满射；(2)  $\forall x, y \in N_1$ ，满足  $\rho_1(x, y) = \rho_2(\phi(x), \phi(y))$ ，则称两个度量空间  $(N_1, \rho_1)$  和  $(N_2, \rho_2)$  是等距同构的。

嵌入：设存在两个度量空间  $(N_1, \rho_1)$  和  $(N_2, \rho_2)$ ，如果  $(N_1, \rho_1)$  和  $(N_2, \rho_2)$  的子空间  $(N_0, \rho_0)$  是等距同构的，则称  $(N_1, \rho_1)$  可以嵌入  $(N_2, \rho_2)$ 。

Takens 嵌入定理：设是  $d$  维的流形，存在映射  $\phi, y, \phi(\phi, y)$ ，分别满足 (1)  $\phi$  是一个光滑的微分同胚；(2)  $y: M \rightarrow R$ ；(3)  $\phi(\phi, y): M \rightarrow R^{2d+1}$ ，其中  $\phi(\phi, y) = \phi(y(x), y(\phi(x)), y(\phi^2(x)), \dots, y(\phi^{2d}(x)))$ ，则称  $\phi(\phi, y)$  是  $M$  到  $R^{2d+1}$  的一个嵌入。

通过嵌入定理可知，对于时间序列  $x_1, x_2, \dots, x_N$ ，可以通过选取恰当的嵌入维数  $m$  和延迟时间  $\tau$  进行空间重构，使得空间里面的每一个点  $y_i$ ，都有

$y_i = [x(i), x(i+\tau), \dots, x(i+(m-1)\tau)]$ ,  $i = 1, 2, \dots, M$ ,  $M$  是重构的空间中相点的个数,  $M = N - (m-1)\tau$ 。空间重构使得原有空间的各种特性都进行的复制, 因此可以通过确定嵌入维数  $m$  和延迟时间  $\tau$  进行空间重构来还原原来时间序列的某些特性。

## 5.2 确定延迟时间

延迟时间  $\tau$  的空间重构过程需要确定的一个重要参数。选的的延迟时间如果过大, 会导致选取的点之间的相关性很弱, 甚至相互独立; 选取的延迟时间过小, 则会导致相关性太强, 信息表达不充分。常用的确定延迟时间的方法主要有自相关法、平均位移法、互信息复自相关法等, 本文选取互信息法进行延迟时间的确定。

### 互信息法

互信息法<sup>[43]</sup>实际上是根据互信息第一次达到最小值时的延迟作为延迟时间。

设:  $X_i (i = 1, 2, \dots, N), Y_j (j = 1, 2, \dots, M)$  是两个离散的随机变量, 则  $X$  的信息熵  $H(X)$

可以表示为  $H(X) = -\sum_{i=1}^N P(x_i) \log P(x_i)$ , 用来衡量事件发生的平均不确定性, 其中  $P(x_i)$  为

事件  $X_i$  的发生概率。设定在事件  $Y$  发生的条件下事件  $X$  的信息熵  $H(X|Y)$  为:

$$H(X|Y) = -\sum_{i=1}^N \sum_{j=1}^M P(y_j) P(x_i | y_j) \log P(x_i | y_j)$$

其中  $P(y_j)$  表示事件  $Y_j$  发生的概率,  $P(x_i | y_j)$  表示在事件  $Y_j$  发生的条件下事件  $X_i$  发生的条件概率。对  $X, Y$  两个事件, 将事件  $X$  因事件  $Y$  的发生而减少的不确定性称之为互信息熵, 符号表示设为  $I(X, Y)$ , 则  $I(X, Y) = H(X) - H(X|Y)$ 。

在此处引入联合信息的概念, 把两个事件作为联合事件, 则可以定义  $X, Y$  的联合信息

为  $H(X, Y)$ ,  $H(X, Y) = -\sum_{i=1}^N \sum_{j=1}^M P(x_i, y_j) \log P(x_i, y_j)$ , 其中  $P(x_i, y_j)$  表示事件  $X_i$  和  $Y_j$  同时发生

的联合概率。此时互信息熵  $H(X, Y)$  又可以表示为:

$$I(X, Y) = H(X) + H(Y) - H(X, Y)$$

在实际运用互信息法进行延迟接触的求解过程中, 对任意时间序列  $\{x_i, i = 1, 2, \dots, N\}$ , 假设延迟时间为  $\tau$ , 时间序列中  $x_i$  出现的概率为  $p(x_i)$ ,  $x_{i+\tau}$  出现的概率为  $p_{i+\tau}$ ,  $x_i$  和  $x_{i+\tau}$  的

联合概率为  $p(x_i, x_{i+\tau})$ ，可以求出  $x_i$  的信息熵  $H(x_i)$ 、 $x_{i+\tau}$  的信息熵  $H(x_{i+\tau})$  以及联合信息熵  $H(x_i, x_{i+\tau})$ ，则对于已知的时间序列，互信息熵  $I(x_i, x_{i+\tau})$  可以表示为延迟时间  $\tau$  的函数，可以通过绘制  $I(x_i, x_{i+\tau})$  和  $\tau$  的函数图像找到第一次使互信息熵达到最小值时对应的  $\tau$ ，这个  $\tau$  就是求解的延迟时间。

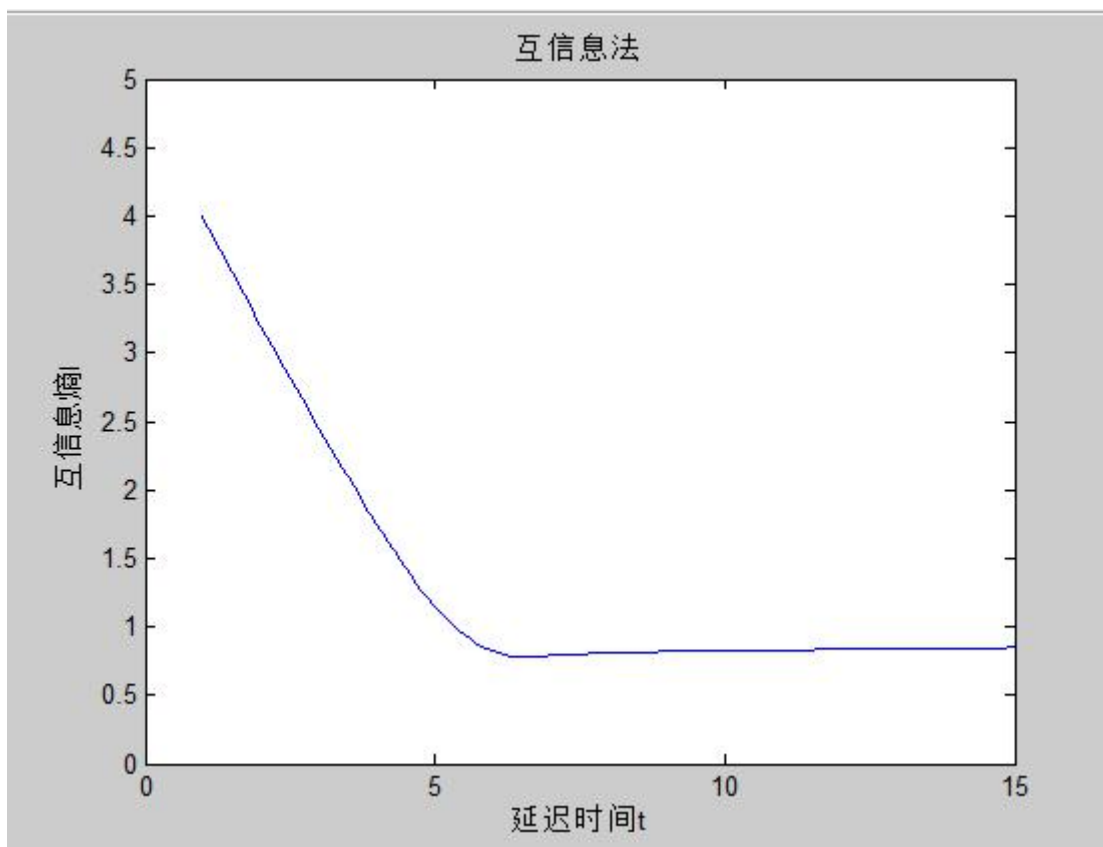


图 3-1 互信息法求延迟时间

实际运用互信息法求解黄金价格的结果如图 3-1，由图中可以看出，互信息熵第一次达到最小值的时候，延迟时间为 7，因此可以判断黄金价格的混沌时间序列在空间重构过程中应选取的延迟时间是 7。

### 5.3 确定嵌入维数

嵌入维数  $m$  是空间重构需要确定的另一个重要参数，根据 Takens 的嵌入定理，只要嵌入维数  $m \geq 2d + 1$  就可以进行空间的重构，但是在实际的过程中需要选取合理的嵌入维数来更好的重构空间。过大的延迟时间会使得相点个数减少，包含的信息不够多，容易产生偏差；过小的嵌入维数会使得嵌入空间无法很好地容纳混沌吸引子，构造的空间和原来混沌序列之间的特性会产生差异。通常用来确定嵌入维数的方法主要有 G-P 算法、伪邻进点算法、CAO 方法等，本文选取 CAO 方法来进行嵌入维数的确定。

## CAO 方法

CAO 方法<sup>[44]</sup>是华裔学者曹良月在为邻进点基础上改进的用于确定最小嵌入维数的一种方法。

设时间序列  $\{x_i, i=1,2,\dots,N\}$  空间重构的点为  $y_i(m)=[x(i),x(i+\tau),\dots,x(i+(m-1)\tau)]$ ,  $i=1,2,\dots,N-(m-1)\tau$ , 其中  $m$  表示嵌入维数,  $\tau$  表示延迟时间,  $y_i(m)$  表示嵌入维数是  $m$  的重构空间中第  $i$  个向量, 定义  $a(i,m)=\frac{\|y_i(m+1)-y_{n(i,m)}(m+1)\|}{\|y_i(m)-y_{n(i,m)}(m)\|}$ , 其中  $y_i(m+1)$  表示嵌入维数为  $m+1$  的重构空间中第  $i$  个向量,  $y_i(m+1)=[x(i),x(i+\tau),\dots,x(i+m\tau)]$ ,  $y_{n(i,m)}$  表示  $m$  维重构的空间中  $y_i(m)$  的最邻近点。

如果嵌入维数是正确的嵌入维数, 那么在重构的  $m$  空间中选取的两个邻近点在重构的  $m+1$  维空间中也会是两个邻近点。如果在重构的  $m+1$  维空间中两个点不是邻近点, 距离很大, 那么就说明  $m$  不是合适的想空间重构维数,  $m$  维空间的两个邻近点成为伪邻近点。在最佳嵌入维数重构的空间中不存在伪邻近点。

通过转化, 在确定嵌入维数时变成了确定是不是所有的点都是邻近点, 此时通过确定阈值来与  $a(i,m)$  进行比较。不同的时间序列对应的阈值应该是不同的, 为了避免阈值不

定带来的问题。引入  $a(i,m)$  的平均值  $E(m)$  作为评价的参考标准,  $E(m)=\frac{1}{N-m\tau}\sum_{i=1}^{N-m\tau}a(i,m)$ ,

此时的  $E(m)$  就只与延迟时间及嵌入维数有关。引入一个新的量  $E_1(m)$ , 令  $E_1(m)=\frac{E(m+1)}{E(m)}$ ,

从公式中可以看出, 在混沌时间序列中, 当嵌入维数达到最小嵌入维数之后, 分子和分母将趋于相等, 因此可以再确定延迟时间的基础上,  $E_1(m)$  将表现出与嵌入维数  $m$  的函数关系, 当达到最小嵌入维数之后,  $E_1(m)$  将趋于稳定, 不再变化。因此当一个时间序列的嵌入维数  $m$  大于  $m_0$  时,  $E_1(m)$  就不在变化了, 那么最小的嵌入维数就是  $m_0$ 。为了区分不同的时间序列类型, 引入  $E_2(m)$

$$E_2(m)=\frac{E^*(m+1)}{E^*(m)}, \text{ 其中 } E^*(m)=\frac{1}{N-m\tau}\sum_{i=1}^{N-m\tau}|x_{(i+m\tau)}-x_{(n(i,m)+m\tau)}|$$

如果一个时间序列是随机的那么  $E^*(m)$  和  $E^*(m+1)$  应当是相等的,  $E_2(m)$  表现恒等于

常数 1. 因此可以用  $E_2(m)$  来区分时间序列的性质。

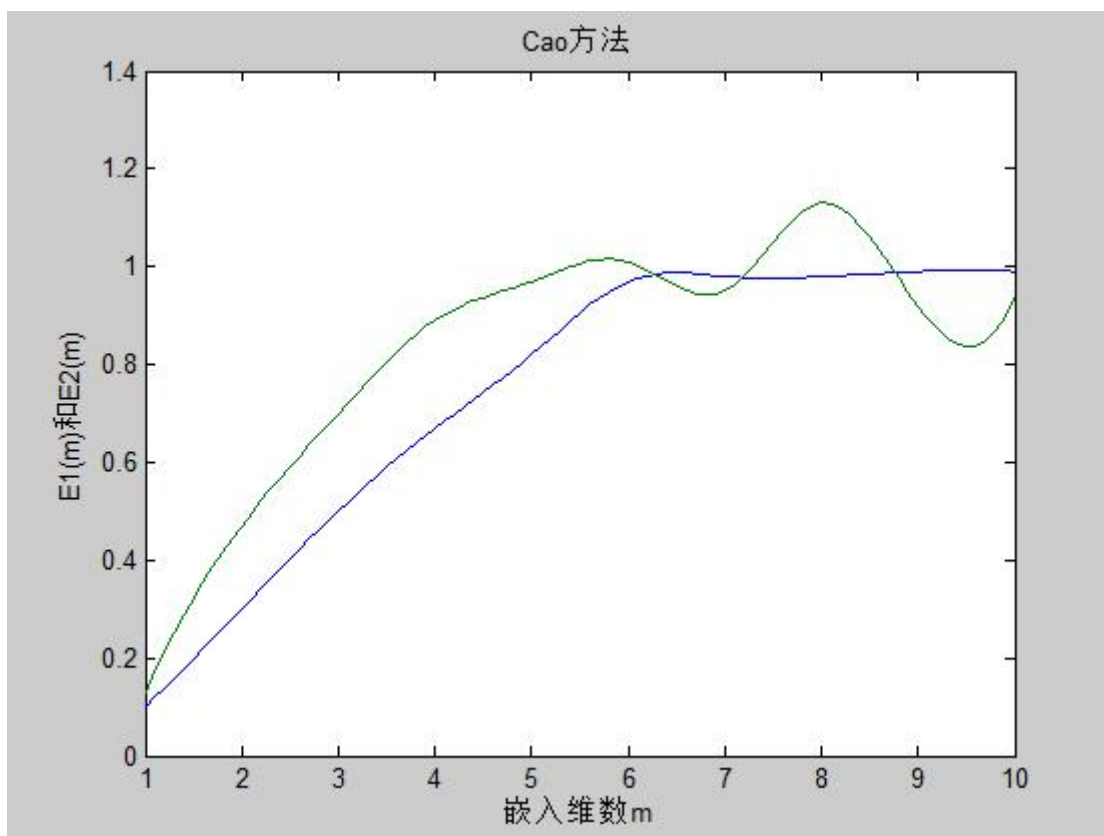


图 3-2 CAO 方法求嵌入维数

在实际运用 CAO 方法对黄金价格求解的结果如图 3-2, 从图中可以看出嵌入维数为 6 的时候  $E_1(m)$  逼近 1, 同时通过  $E_2(m)$  也可以判断是来时间序列并不是随机的。综合判断, 得出黄金价格空间重构确定的嵌入维数应当是 6。

## 5.4 本章小结

本章属于全文的基础章节, 主要是对混沌时间序列进行空间重构。通过重构空间, 提取延迟时间和嵌入维数, 为后面的章节提供相关的参数。介绍了空间重构的原理即 Takens 嵌入定理, 空间中求解延迟时间的互信息法, 空间重构中确定最小嵌入维数的 CAO 方法, 并对黄金价格时间序列求出延迟时间和嵌入维数。将嵌入维数作为神经网络模型中的神经元个数参考值。

## 第六章、神经网络模型预测黄金价格

### 6.1 数据的预处理

通过参考前人对黄金价格的实证分析，总结出黄金价格主要与人民币美元汇率、大盘指数、通货膨胀率、原油价格等因素有关。本文做的是短期的黄金价格预测，通货膨胀率在我国是每个月公布的阶段性数据，因此在分析的时候暂不考虑。黄金价格选自中国黄金市场 2015 年的数据，人民币美元汇率取自中国人民银行公布的汇率中间价，股票指数取自上证指数，原油价格选取的是美国市场的轻质原油价格。同时考虑到前人对神经网络模型的预测研究多是基于自身特点的数据进行分析的，因此本文还选取前一日收盘价、前一日成交量、当日开盘作为影响当日收盘价的相关影响因素。全年中国黄金市场共开盘 244 天，所以数据总共有 244 组。在数据预处理之前，为了直观反映出 2015 年黄金价格走势，这里先绘出 2015 年黄金价格走势图，全年黄金价格走势如图 6-1，从图中可以看出 2015 年黄金价格整体是下跌的走势其中在图中第 50 期、130 期、200 期左右的位置有明显的突然下跌状况。

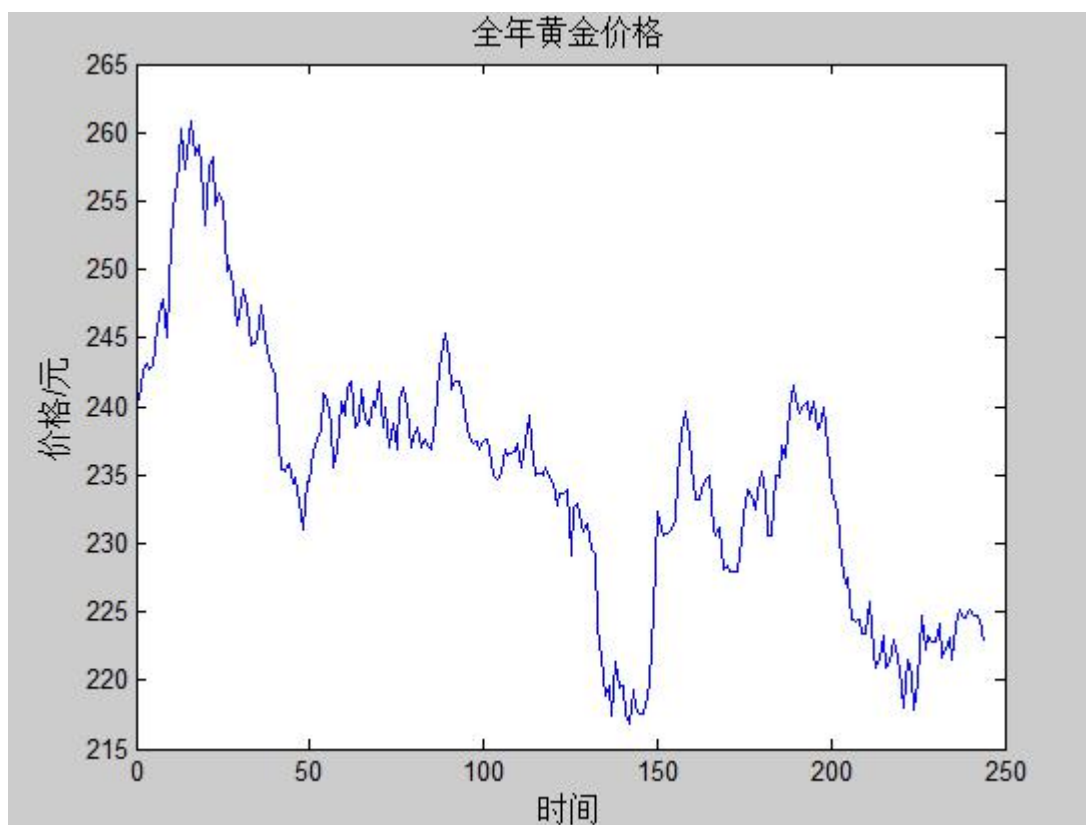


图 6-1 2015 年黄金价格走势

在本文中数据预处理的方法主要是对数据进行归一化处理，因为包括成交量、交易价格、人民币美元汇率等数据之间的标量是不相同的，并且数据数据差别很大，在此，对数

据进行归一化处理。把所有的数据都统一归一到[0,1]之间。

设时间序列  $x_i, i=1,2,\dots,N$ ，令  $x_{\max}$  表示序列中的最大值， $x_{\min}$  表示序列中的最小值，对于序列中的任何一点  $i$ ，归一化之后的数值  $y_i$  表示为：

$$y_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

通过归一化处理，可以将有量纲的表达式，经过变换，化为无量纲的表达式。有利于将数据统一处理，减少预测过程中因为数据量纲对结果产生的影响。

部分数据的预处理结果如表 6-1

表 6-1 归一化后的数据

日期	当日开盘价	前一日收盘价	前一日成交量	人民币美元汇率	上证指数	轻质原油价格
2015/1/5	0.5941	0.5262	0.5316	0.0438	0.1278	0.6860
2015/1/6	0.6018	0.5387	0.3335	0.0459	0.1606	0.5056
2015/1/7	0.6319	0.5831	0.5932	0.0493	0.1587	0.5342
2015/1/8	0.6499	0.5991	0.4768	0.0578	0.1794	0.5398
2015/1/9	0.6369	0.5888	0.5316	0.0563	0.1361	0.5231
2015/1/12	0.6536	0.5929	0.4634	0.0399	0.1276	0.4192
2015/1/13	0.6917	0.6583	0.4707	0.0301	0.1118	0.4343
2015/1/14	0.7415	0.6927	0.4200	0.0327	0.1203	0.4328
2015/1/15	0.7316	0.7039	0.4167	0.0296	0.1120	0.4380
2015/1/16	0.7740	0.6424	0.5784	0.0283	0.1664	0.5233

## 6.2 相关性分析

进行相关性分析的目的是通过得到的数据，验证前人对黄金价格实证分析结论在本文数据中的可靠性，并且通过分析结果，为神经网络模型预测时选择输入变量提供依据。这里对相关数据和黄金收盘价格进行相关性分析，选取相关系数作为指标进行考量。设

$$x_i, y_j, i, j = 1, 2, \dots, n, \text{ 为两个序列, } x, y \text{ 的相关系数为 } r, \text{ 则 } r = \frac{\sum_{i,j=1}^n (x_i - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 * \sum_{j=1}^n (y_j - \bar{y})^2}}。$$

相关性分析结果如下表 6-2 的结果:

表 6-2 相关性分析表

相关指标	开盘价	成交量	收盘价	人民币美元汇率	上证指数	轻质原油价格
相关系数 (R)	0.9782	0.0787	0.9780	-0.4827	-0.0554	0.3767
P 值	0.00	0.22	0.00	0.00	0.39	0.00

通过相关性分析的结果可以看出,人民币美元汇率、上证指数和黄金价格呈现负相关,轻质原油价格、开盘价、成交量、成交价及收盘价与黄金价格呈现正相关。其中汇率、股票指数、原油价格对黄金价格的影响方向和文中第二章部分总结的黄金价格的其他市场因素中相关指标的影响方向一致。在相关性检验上,成交量和上证指数的 P 值都明显大于 0.05,因此可以认为上证指数、成交量与黄金价格的相关性并不是统计上明显的,在选取神经网络模型输入变量时,不考虑上证指数、成交量两个相关性不显著的因素。

## 6.3 BP 神经网络预测黄金价格

### 6.3.1 BP 神经网络算法设计步骤

BP 神经网络算法实验过程中的基本步骤如下图 6-2

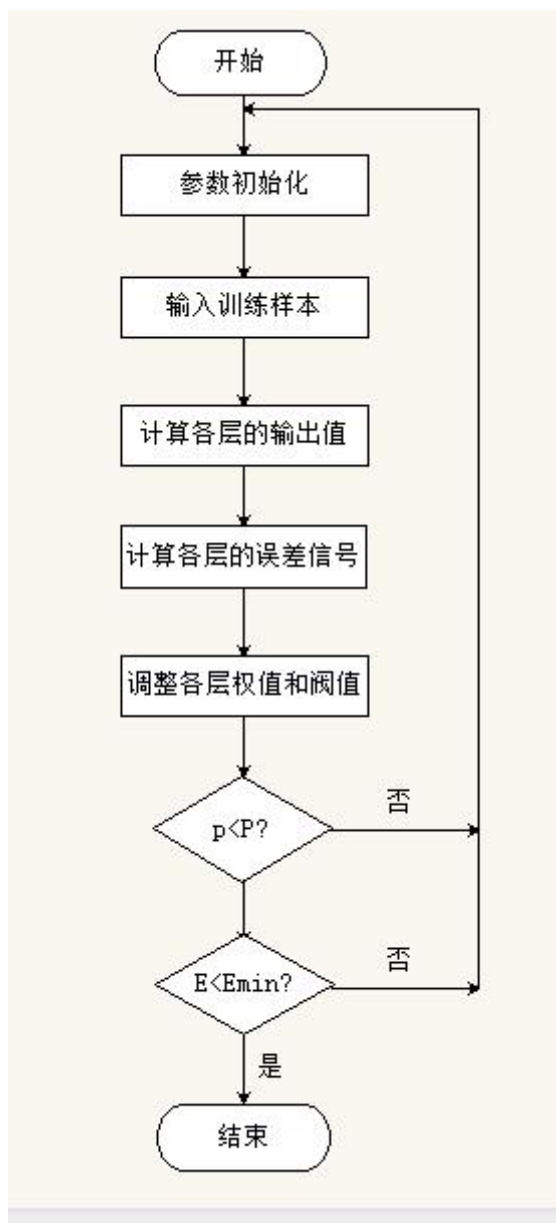


图 6-2 BP 神经网络学习算法流程图

- 1、参数初始化，随机初始化网络中的权值矩阵和阈值，使初始化训练时  $E = 0$ ，设置样本数  $P$ 、最小训练误差  $E_{\min}$  和学习率  $\eta, 0 < \eta < 1$
- 2、输入训练样本，计算出网络中各层的输出值
- 3、根据公式计算网络中输出值和期望输出值之间的误差，并计算总体输出误差
- 4、计算各层的误差信号  $\delta$
- 5、根据设定公式调整权值
- 6、检查训练过程是否完成，及  $P$  个样本是否都进行训练，如果没有，返回第 2 部直至全部训练完成；如果完成，则进行下一步
- 7、检查网络的总体输出误差是否小于既定要求  $E_{\min}$ ，如果满足，则训练结束；否则则重复

进行，直至满足。

### 6.3.2 BP 神经网络的参数选取及预测流程

BP 神经网络在应用过程中主要可以通过主要从网络的层数、神经元的个数、初始权值和阈值的选取以及学习速率四个方面考虑。

神经网络层数的确定指的是隐含层个数的确定，原则上一个隐含层就可以对输入通过映射向期望的输出结果进行逼近，但是很多时候只有一个隐含层的神经网络模型对结果的逼近并不是最优的。本文在设计神经网络隐含层个数的时候，开始设置一个隐含层，然后通过增加隐含层分别得到预测的输出结果，将预测的输出结果和实际预留的真实输出结果进行比较，选择预测结果明显更加准确的神经网络模型的隐含层个数作为网络设计层数参考值。最终选择的隐含层数是一层。

网络神经元的个数选择是一个很难找到最优值的结果。如果神经元个数较少，会导致模型学习能力不足，即使增加层数也无法很好的解决问题，容易陷入局部极小值甚至得不到稳定的结果；如果神经元节点过多，又会导致出现过度拟合，使不必要的信息都拟合到模型中去。模型最终最小嵌入维数作为模型神经元个数。

对于初始的权值和阈值的选取，因为考虑到没有先验信息作为参考，阈值和权值由系统随机形成。

学习效率在 BP 神经网络模型中是一经确定不再改变的，如果选择的效率过小，每一次调整幅度就会很小，最终得到收敛结果需要的时间就会很长；如果效率过大，可能因为调整幅度过大而错过误差最小值的结果。在实际应用的时候一般性选择较小的学习速率来保证结果的准确性，取值范围一般在 0.01-0.8 之间，这里因为数据量并不大，为了保证结果的准确性，选择的学习效率为 0.01。

具体的 BP 神经网络实验设计流程图如下图 6-3

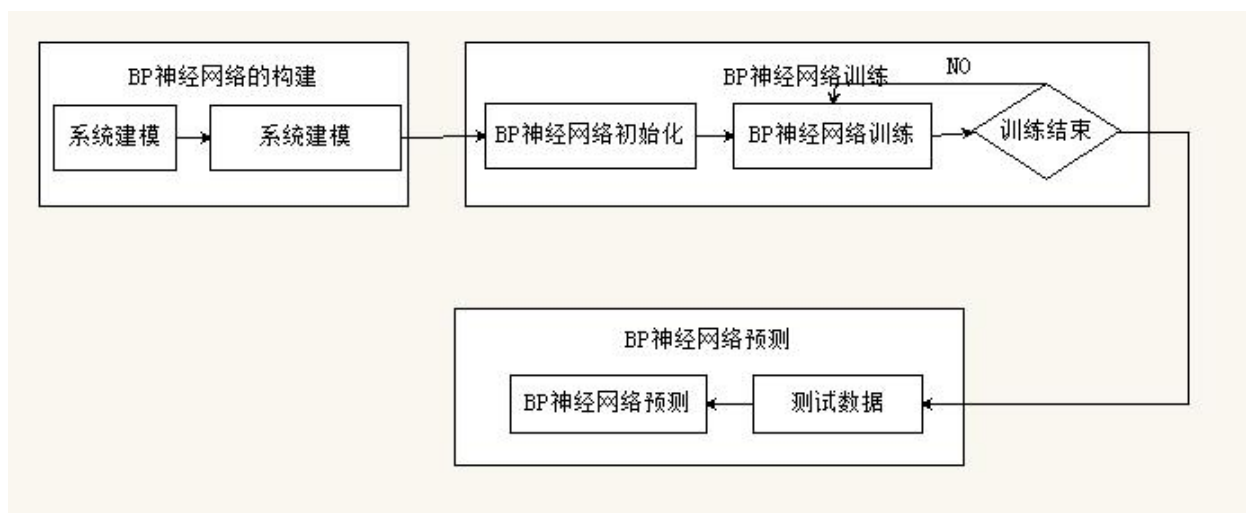


图 6-3 BP 神经网络实验设计流程图

### 6.3.3 BP 神经网络预测结果

在实际模拟过程中考虑到数据量以及数据波动对结果的影响，全年共有 244 组数据，并且数据波动比较大在选择模型输入量个数的时候，分别从 0 时刻、50 时刻、100 时刻和 150 时刻进行截取选定时刻后面的数据作为输入变量进行 BP 神经网络的模拟，发现从时刻 0 开始选取的数据，在 BP 网络中模拟预测的效果最好，同时对于本文选取的具有相关性的四个变量，在选取的输入变量是当天开盘价和人民币美元汇率的时候，模型的预测效果最好。黄金价格运用 Matlab 进行黄金价格预测结果如下图 6-4 和图 6-5。

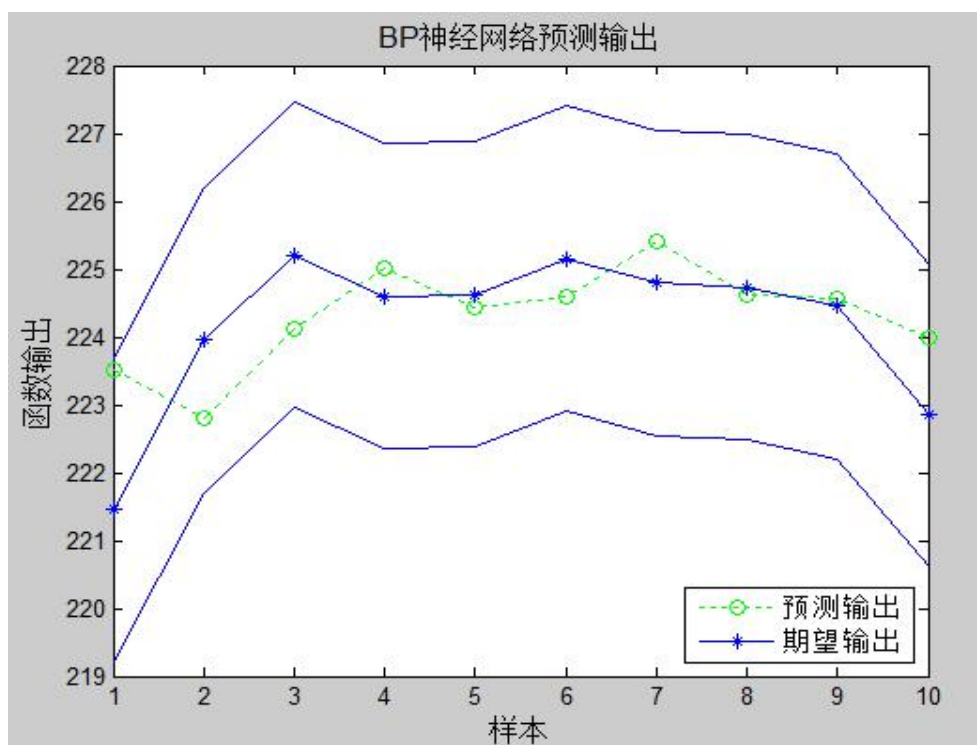


图 6-4 BP 神经网络预测 10 期结果图

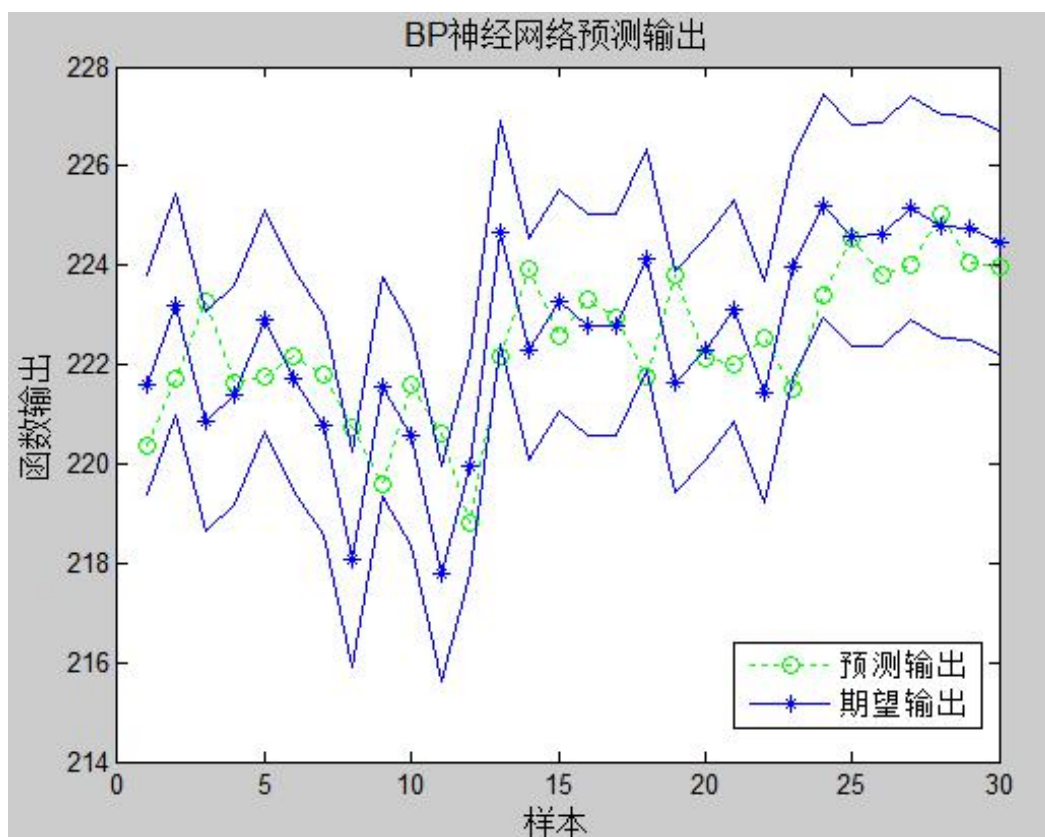


图 6-5 BP 神经网络预测 30 期结果图

图 6-4 中是用 15 年黄金价格模拟的 BP 神经网络，选取前 234 组数据作为训练数据，留用 10 个数据作为预测数据的预测结果图。图 6-5 中是用 15 年黄金价格模拟的 BP 神经网络，选取前 214 组数据作为训练数据，留用 30 个数据作为预测数据的预测结果图。图中上线两条线分别为实际黄金价格上下波动 1% 时价格区间，从图中可以看出整体的拟合效果不错，预测价格都在真实价格 1% 的区间之内。可以用 BP 神经网络来预测黄金价格。

预测的结果误差图如下图 6-6 和图 6-7:

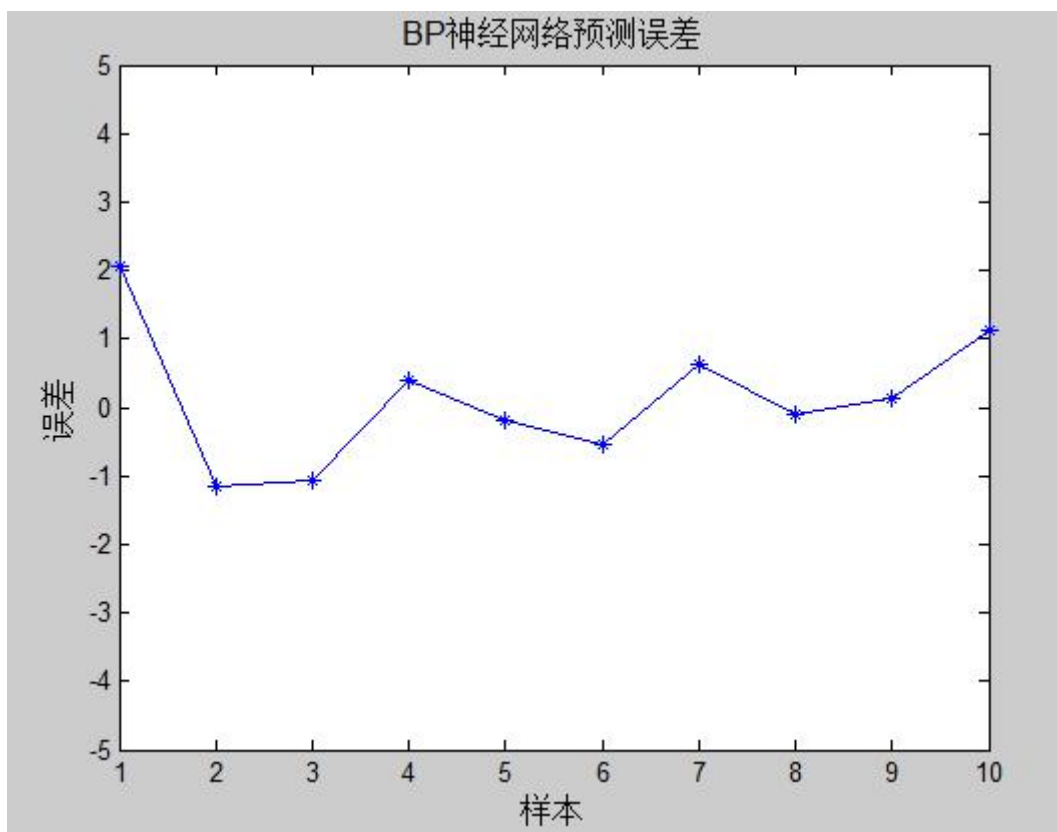


图 6-6 BP 神经网络预测 10 期误差图

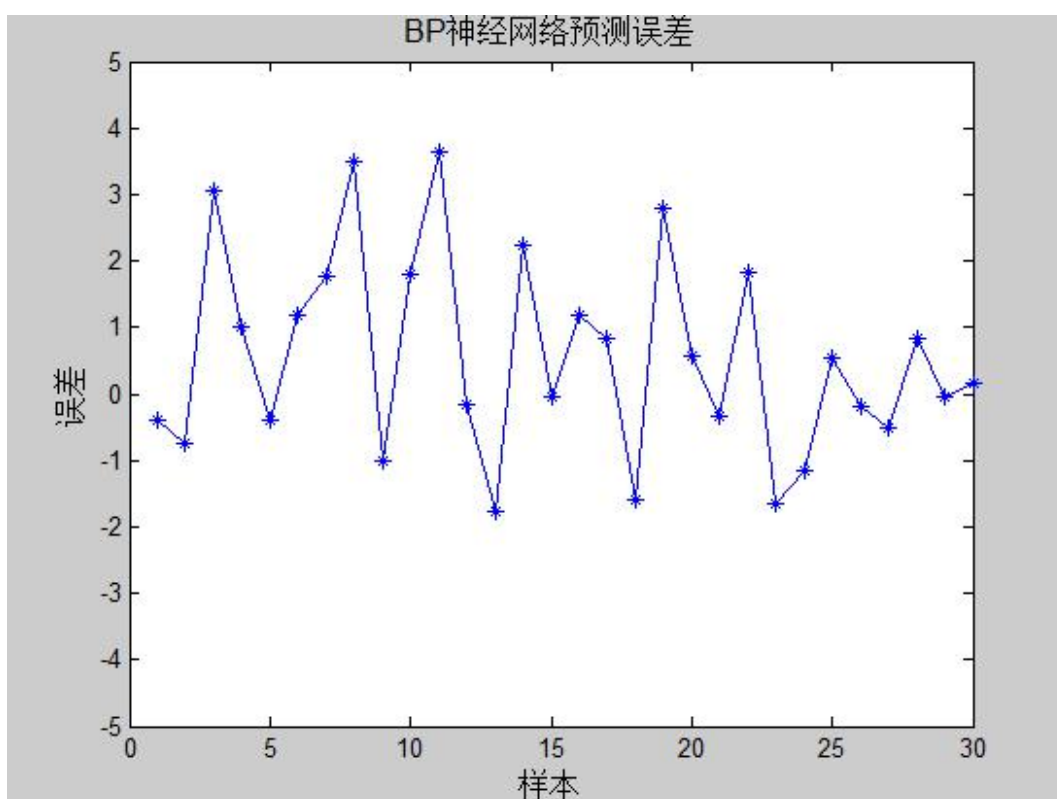


图 6-7 BP 神经网络预测 30 期误差图

同时画出如上图的 BP 神经网络预测模型的结果误差图，从图中可以看出模型拟合的

整体误差并不是很大，但随着预测期数的增多，模型预测的误差开始变大，说明在用 BP 神经网络预测黄金价格的时候，预测的时期越长，误差就会越大，因此短期预测效果会更好一点。计算整体 10 期预测的误差平方和为 9.484，整体 30 期预测的误差平方和为 75.096，为了比较，求出平均每期误差平方，预测期为 10 的，平均每期误差平方和为 0.9484，预测期为 30 的，平均每期误差平方和为 2.5023。通过比较均方误差，可以看出预测期短的，预测效果明显好于预测期更长的。

## 6.4 RBF 神经网络模型设计及预测

RBF 神经网络的模型设计思路同样类似与 BP 神经网络模型的设计方法。首先还是选择样本数据，包括样本的输入和输出数据，并且确定隐含层神经元的个数，这里隐含层神经元的个数确定为 6，即空间重构过程中确定的嵌入维数。在运用神经网络模型预测的过程中，通过训练部分的样本会得到一个神经网络的模型，确定模型的参数。实际预测的时候，会将预测部分输入逐个带入带入模型当中，并且把上一个样本数据的输出值作为下一个待预测样本的输入值不断循环迭代进行预测。整个的数据处理、模型的参数求解和预测过程如下图 6-8 结构。

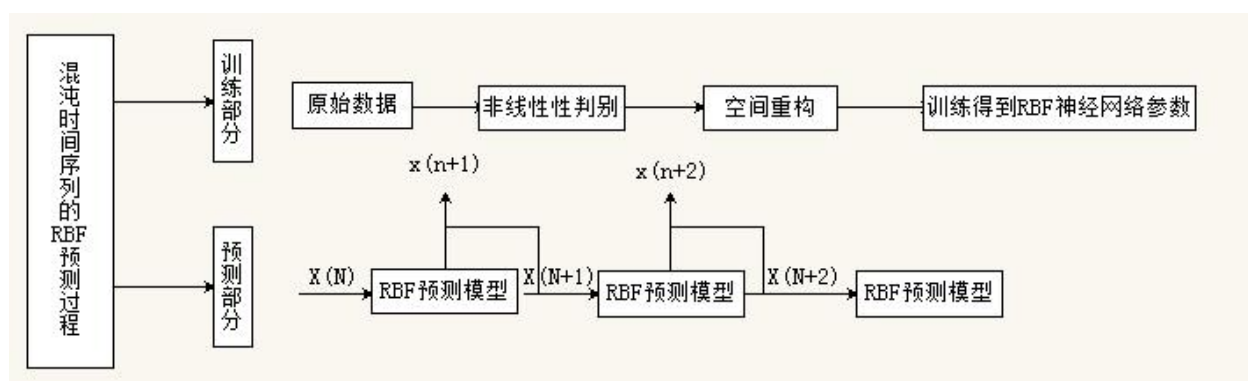


图 6-8 RBF 神经网络实验设计图

考虑到全年共有 234 组数据，并且数据波动比较大在选择模型输入量个数的时，分别从图 4-8 所示的 0 时刻、50 时刻、100 时刻和 150 时刻进行截取选定时刻后面的数据为输入变量进行 RBF 神经网络的模拟，发现从时刻 100 开始选取的数据，在 RBF 网络中模拟预测的效果最好，同时对于输入数据的选取，也是选取当天开盘价和人民币美元汇率的时候模型预测的均方误差最小，模型预测效果最好。运用 Matlab 进行黄金价格预测结果如下图 6-9 和图 6-10。

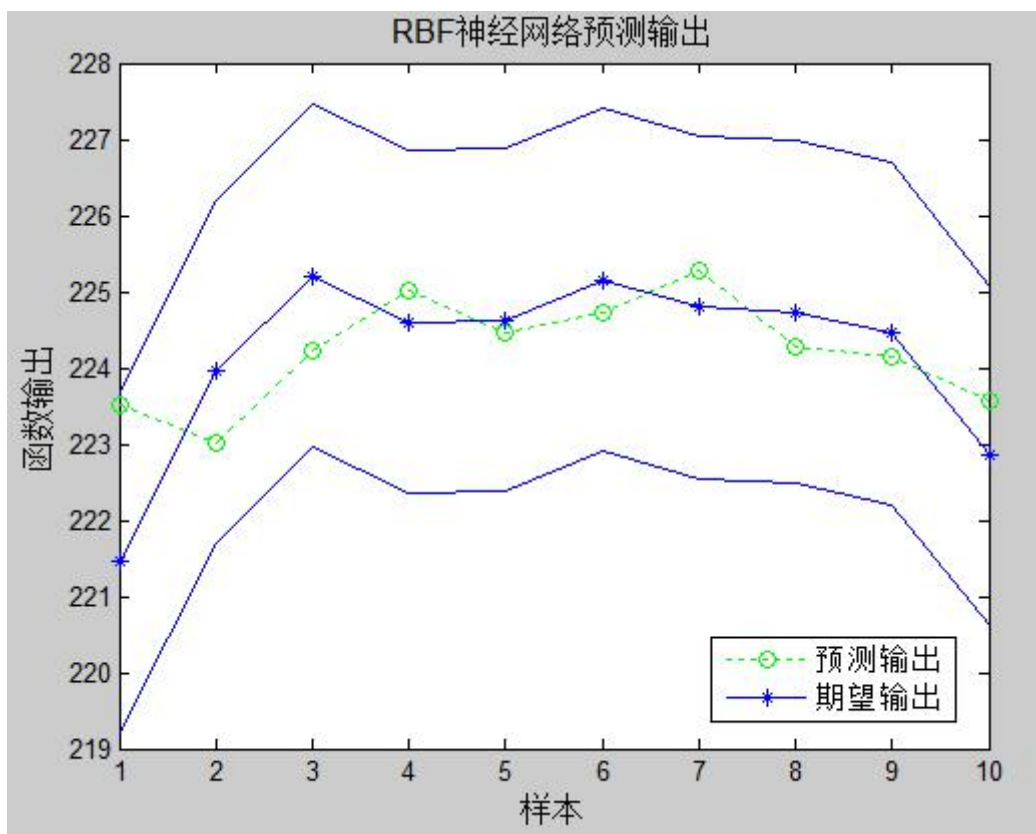


图 6-9 RBF 神经网络预测 10 期结果图

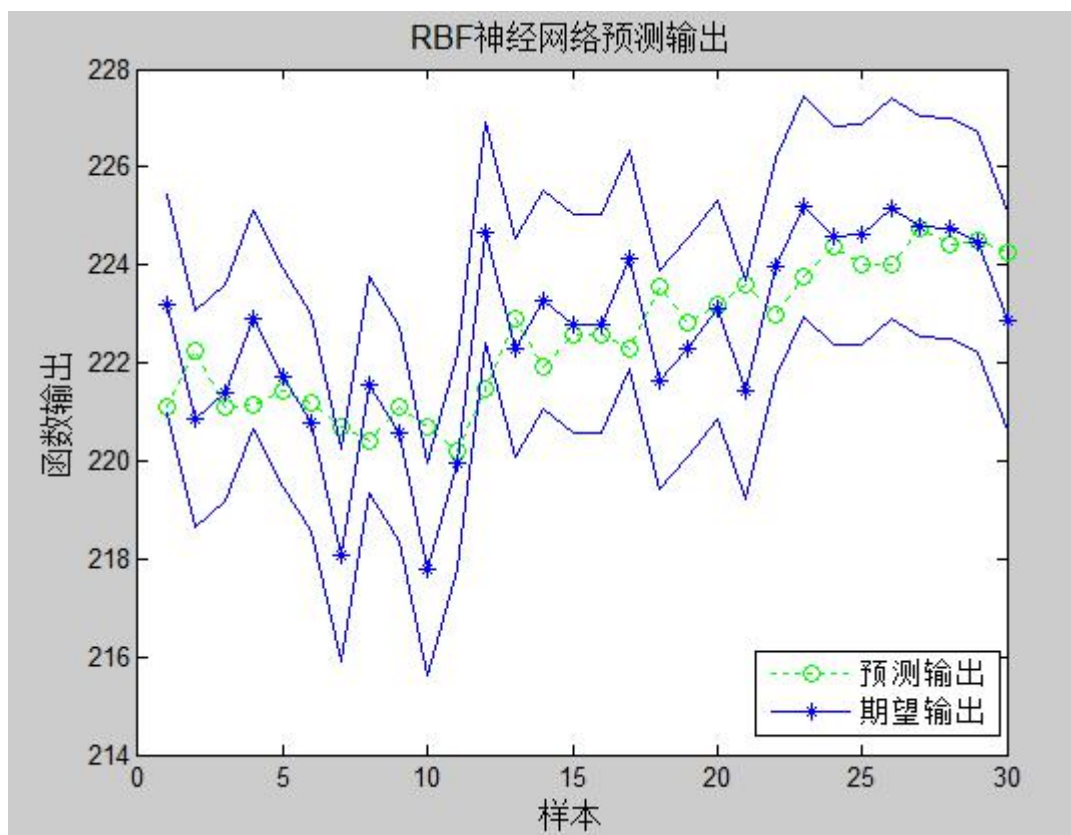


图 6-10 RBF 神经网络预测 30 期结果图

图 6-9 中是用 15 年黄金价格模拟的 RBF 神经网络, 选取 15 年黄金价格数据的第 100

组开始作为训练数据，留用 10 个数据作为预测数据的预测结果图。图 6-10 中是选取 15 年黄金价格数据的第 100 组开始作为训练数据，留用 30 个数据作为预测数据的预测结果图。节点个数为 6 个，初始权值和阈值选取方法和 BP 神经网络相同，因为对先验信息不了解，也不考虑正则化影响因素。得到如图模拟结果，图中上线两条线分别为实际黄金价格上下波动 1% 时价格区间，从图中可以看出整体的拟合效果不错，预测价格都在真实价格 1% 的区间之内。可以用 RBF 神经网络来预测黄金价格。

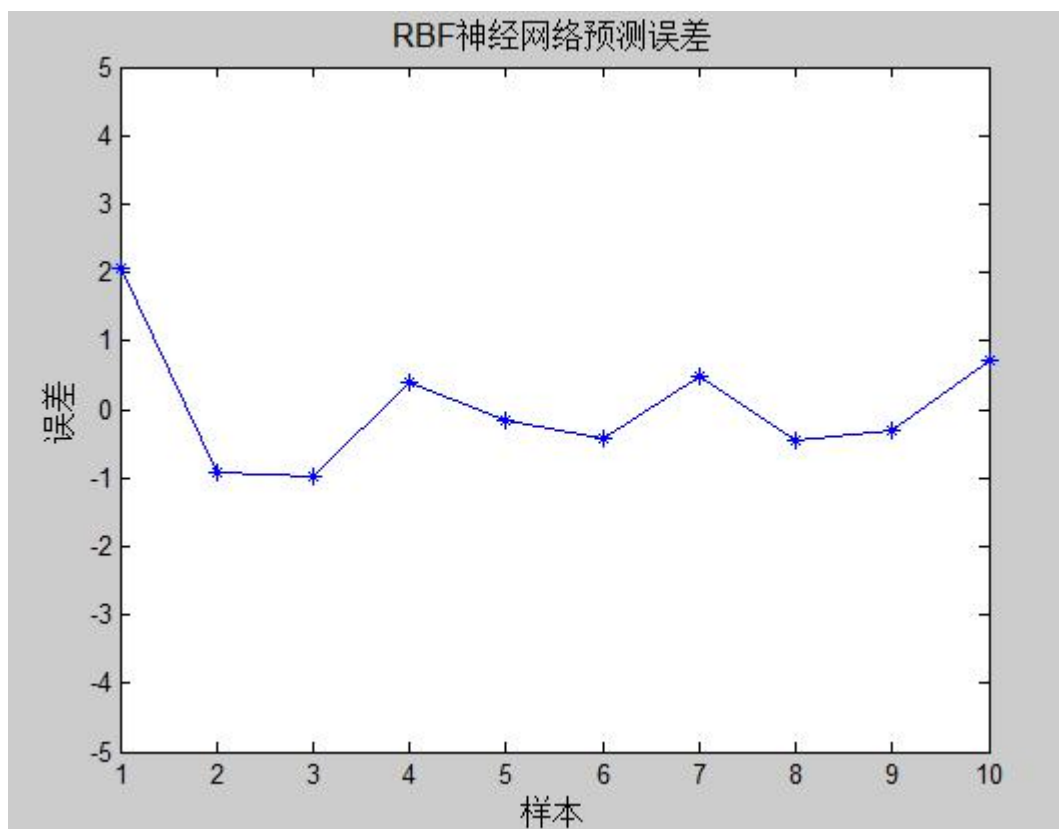


图 6-11 RBF 神经网络预测 10 期误差图

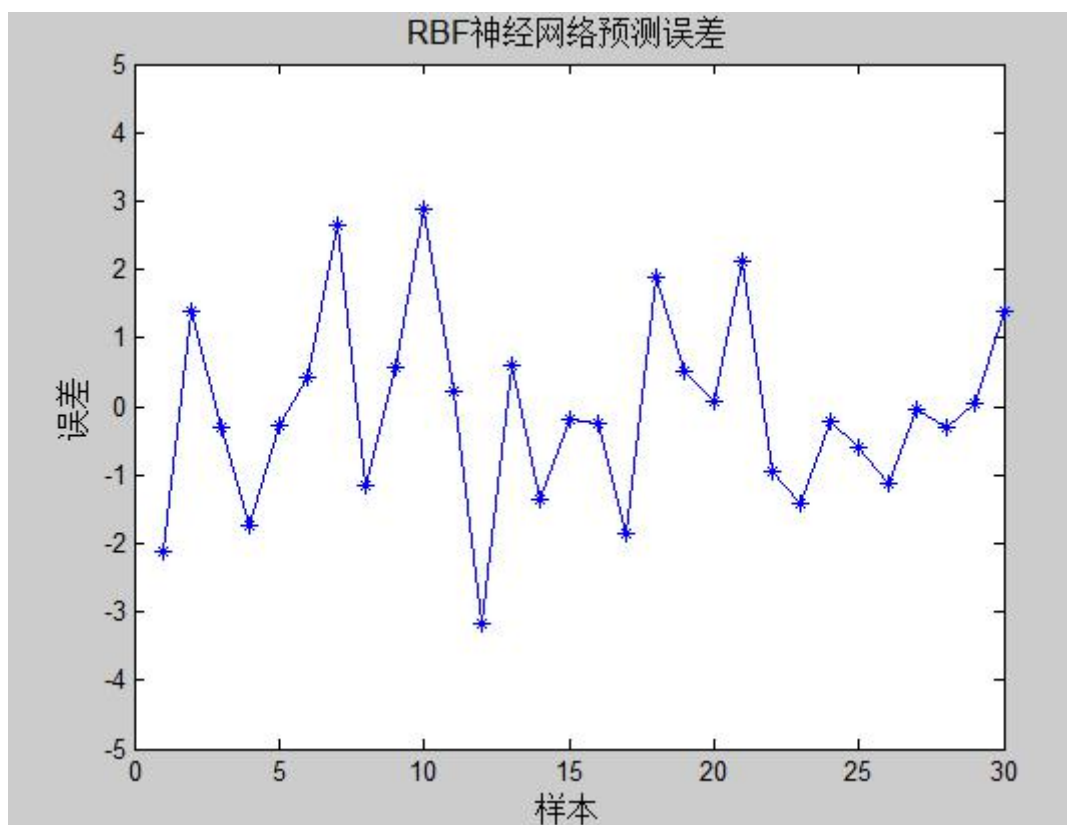


图 6-12 RBF 神经网络预测 30 期误差图

同时画出如上图的 RBF 神经网络预测模型的结果误差图,从图中可以看出模型拟合的整体误差并不是很大,但随着预测期数的增多,模型预测的误差开始变大,说明在用 RBF 神经网络预测黄金价格的时候,预测的时期越长,误差就会越大,因此短期预测效果会更好一点。计算整体 10 期预测的误差平方和为 7.504,整体 30 期预测的误差平方和为 57.594,为了比较,求出平均方误差,预测期为 10 的,均方误差为 0.7504,预测期为 30 的,均方误差为 1.9198。通过比较平均每期的误差平方,可以看出预测期短的,预测效果明显好于预测期更长的。

## 6.5 BP 神经网络和 RBF 神经网络预测结果比较

通过分别运用 BP 神经网络和 RBF 神经网络对混沌的黄金价格进行预测可以看出,在本文的预测中,两种神经网络模型都可以对结果进行很好的预测。满足神经网络模型非线性拟合效果优良的特点。从模型预测的精准度来看,RBF 神经网络模型预测结果,在 10 期预测的均方误差上仅为 0.7504 小于 BP 神经网络模型的 0.9484,在 30 期的预测上均方误差为 1.9198 同样小于 BP 神经网络模型的 2.5023。并且 RBF 神经网络模型运用的模拟数据更少,在运用更少的数据基础上,得到比 BP 神经网络模型更精准的预测结果,因此认为在对 2015 年黄金价格预测中,RBF 神经网络模型比 BP 神经网络模型有更好的预测效果。

## 6.6 本章小结

本章是本文的最重要的一个部分。主要内容是运用 BP 神经网络模型和 RBF 神经网络模型对黄金价格进行预测。首先开始通过两个小节简单介绍了本文选取数据的原因以及对选取数据进行的预处理。然后分别运用 BP 神经网络模型和 RBF 神经网络模型对黄金价格进行 10 期和 30 期的预测，通过结果发现两种神经网络模型都可以很好的对黄金价格进行预测。最后比较两种模型的预测结果，发现 RBF 神经网络模型的预测结果要比 BP 神经网络模型的预测结果更加精准。

## 第七章、总结与展望

本文主要是通过 BP 神经网络和 RBF 神经网络对中国黄金市场上的黄金价格进行研究。通过分析黄金对国家和消费者的影响说明研究和预测黄金价格具有重要意义。在具体对黄金价格的研究中，进行了价格序列的非线性判别、空间重构、神经网络预测。判别过程中通过功率谱法和 Lyapunov 指数法分别对序列进行定性和定量判别，显示黄金价格是混沌的，从而说明黄金价格序列是非线性的；在空间重构中，分别运用互信息法和 CAO 方法确定延迟时间和最小嵌入维数，并把最小嵌入维数作为神经网络模型的神经元个数；在最后进行价格预测时，分析相关因素和黄金价格的相关性，运用 BP 神经网络和 RBF 神经网络对黄金价格进行预测并和搜集到的真实值进行对比，结果显示两种神经网络模型都对短期结果进行了比较有效的预测，而且 RBF 网络模型对结果的预测误差更小一些。

本文虽然运用 BP 和 RBF 神经网络成功的预测的短期的黄金价格，但是在当价格遭遇黑天鹅现象时，仍然不能进行很好预测。未来在对价格突然变化、短期剧烈波动的情况下的研究有待进一步进行。

## 参考文献

- [1] 汪凯. 国际黄金价格影响因素分时段实证分析[J]. 中国外资, 2012, 08: 43-244.
- [2] 万博, 张华硕. 后金融危机时代国际黄金价格影响因素实证分析[J]. 科技和产业, 2014, 08: 168-172.
- [3] 吴瑶明, 杨帆. 金融危机前后黄金价格影响因素的动态分析[J]. 黄金, 2014, 08: 10-14.
- [4] 董益盈. 货币流动性对黄金价格影响的实证研究[J]. 世界经济情况, 2008, 12: 23-29.
- [5] 新浪财. 2015年中国黄金消费量985吨 同步增长3.66%[OL]. [2016-02-03]. <http://finance.sina.com.cn/money/nmetal/2016-02-03/doc-ifxnzanh0643294.shtml>.
- [6] 新华社. 2015年产金450吨 我国连续九年全球黄金产量第一[OL]. [2016-02-04]. [http://www.cs.com.cn/xwzx/jr/201602/t20160204\\_4900401.html](http://www.cs.com.cn/xwzx/jr/201602/t20160204_4900401.html).
- [7] Jastram, Roy W. The Golden Constant: The English and American Experience[J]. *General Information*, 1977, 31: 168-169.
- [8] J Beckmann, R Czudaj. Gold as an inflation hedge in a time-varying coefficient framework[J]. *North American Journal of Economics & Finance*, 2013, 24(1): 208-222.
- [9] JA Batten, C Ciner, BM Lucey. On the economic determinants of the gold-inflation relation[J]. *Resources Policy*, 2014, 41(3): 101-108.
- [10] PK Narayan, S Narayan, X Zheng. Gold and oil futures markets: Are markets efficient?[J]. *Applied Energy*, 2010, 87(10): 3299-3303.
- [11] YS Wang, YL Chueh. Dynamic transmission effects between the interest rate, the US dollar, and gold and crude oil prices[J]. *Economic Modeling*, 2013, 30(1): 792-798.
- [12] HF Chang, LC Huang, MC Chin. Interactive relationships between crude oil prices, gold prices, and the NT-US dollar exchange rate—A Taiwan study[J]. *Energy Policy*, 2013, 63(63): 441-448.
- [13] G Grudnitski, L Osburn. Forecast S&P and gold futures prices: An application of neural networks[J]. *Journal of Futures Markets*, 2006, 13(6): 631-643.
- [14] Z Ismail, A Yahya, A Shabri. Forecasting gold price using multiple linear regression method[J]. *American Journal of Applied Sciences*, 2009, 6(8): 1509-1514.
- [15] A Sharma, S Baby. Forecasting gold prices using time series analysis[R]. *Second International Conference on Business Analytic & Intelligence*, 2014.
- [16] Gold Silver World. Gold Price 2015: Forecast and prediction[OL]. [2014-12-24]. <http://>

goldsilverworlds.com/price/gold-price-2015-forecasts-predictions/.

[17] Financial Forecast Center. Gold price forecast[OL]. [2016-4-6]. <http://www.forecasts.org/gold.htm>.

[18] 杨柳勇, 史震涛. 黄金价格的长期决定因素分析[J]. 统计研究, 2004, 6: 21-24.

[19] 刘曙光, 胡再勇. 黄金价格的长期决定因素稳定性分析[J]. 世界经济研究, 2008, 2: 35-41.

[20] 付丹, 梅雪, 张晖. 黄金价格与通货膨胀相关性的实证分析[J]. 黄金, 2009, 30(1): 4-7.

[21] 杨叶. 黄金价格和石油价格的联动分析[J]. 黄金, 2007, 28(2): 4-7.

[22] 胡乃联, 宋鑫. 自适应过滤模型在黄金价格预测中的应用[J]. 黄金, 1999, 5: 53-54.

[23] 林雨, 孔刘柳, 刘培. 基于ARFIMA模型的黄金价格预测[J]. 南开大学学报: 社会科学版, 2010, 11(1): 36-38.

[24] 涂小龙. 基于信息修正GM(1,1)模型的黄金价格行情预测[J]. 有色金属, 2015, 64(3): 11-14.

[25] 祝合良. 黄金市场的发展研究[J]. 中国市场, 2012, 33: 45-51.

[26] 靳湘云. 从黄金企业生产成本看金价之底[J]. 中国贵金属, 2013, 6: 56-60.

[27] 乔俊飞, 韩红桂. 前馈神经网络分析与设计[M]. 北京: 科学出版社, 2013.

[28] 谭文, 王耀南. 混沌系统的模糊神经网络控制理论与方法[M]. 北京: 科学出版社, 2008.

[29] 邢红杰, 哈明虎. 前馈神经网络及其应用[M]. 北京: 科学出版, 2013.

[30] 韩敏. 人工神经网络基础[M]. 大连: 大连理工大学出版社, 2014.

[31] 詹姆斯. 心理学原理[M]. 田平, 译. 北京: 中国城市出版社, 2010.

[32] McCulloch W, Pitts W. A logical calculus of the ideas immanent in nervous activity[J]. Bulletin of Mathematical Biophysics, 1943, 5(4): 115-133.

[33] Hopfield J J. Neural networks and physical systems with emergent collective computational abilities[J]. Proceedings of the National Academy of Science, 1982, 79(8): 2554-2558.

[34] R Setiono. Extracting M-of-N rules from trained neural networks[J]. IEEE Transactions on Neural Networks, 2000, 11(2): 512-519.

[35] Werbos P J. Beyond regression: new tools for prediction and analysis in the behavioral

sciences[D]. Ph.D dissertation Harvard University, 1974.

[36] DE Rumelhart, GE Hinton, RJ Williams. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.

[37] DS Broomhead, D Lowe. Radial basis functions, multi-variable functional interpolation and adaptive networks[J]. Advances in Neural Information Processing Systems, 1988, 4148: 728-734.

[38] Moody J, Darken C J. Fast learning in networks of locally-tuned processing units[J]. Neural Computation, 1989, 1(2): 281-294.

[39] Tikhonov A N. Solution of incorrectly formulated problems and the regularization method[J]. Soviet Mathematics-Doklady, 1963, 4(4): 1035-1038

[40] Li T Y, Yorke J A. Period three implies chaos[J]. America Mathematics Monthly, 1975, 82: 985-992.

[41] A Wolf, JB Swift, HL Swinney. Determining Lyapunov exponents from a time series[J]. Physica D, 1985, 16: 285-317.

[42] Floris Takens. Detecting strange attractors in turbulence[M]. Springer Berlin Heidelberg, 1981, 898:366-381.

[43] AM Fraser and HL Swinney. Independent coordinates for strange attractors from mutual information[J] Physical Review A, 1986, 33(2): 1134-1140.

[44] Cao LY. Practical method for determining the minimum embedding dimension of a scalar time series[J]. Physica D Nonlinear Phenomena, 1997, 110:43-50.

[45] 王小川, 史峰, 郁磊, 李洋. MATLAB 神经网络 43 个案例分析[M]. 北京: 北京航空航天大学出版社, 2013.

## 致谢

两年的研究生生活已经接近尾声，在暨南大学读研究生的这两年的时间里，相信将会是人生中最难忘的两年。这两年，收获了很多，也成长了很多。在即将毕业的日子里，感谢这两年给与我帮助的每一个人。

首先，要感谢我的导师范旭乾老师。在这两年的时间里，范老师给与我悉心的教导，教授我知识，同时对我完成毕业论文提供指导。

然后，要感谢数学系的其他老师们。这两年学习生活的离不开数学系每一位老师的帮助，在此感谢参与我们教学的每一位老师。

最后，要感谢我身边的同学们。两年的学习生活，离不开身边的每一个朋友和同学，你们让我这两年的生活更加丰富多彩，在此表示感谢。

## 阅读此文的还阅读了:

- [1. RBF网络在电池容量预测中的应用](#)
- [2. 基于RBF-BP神经网络的管道腐蚀类型预测](#)
- [3. BP神经网络模型在黄金期货价格预测中的实证分析](#)
- [4. 基于BP神经网络的居民消费价格预测研究](#)
- [5. 时间序列在黄金价格预测中的应用](#)
- [6. 统计学在黄金价格分析预测上的应用](#)
- [7. BP神经网络与RBF神经网络在股价预测中的应用](#)
- [8. 基于BP神经网络的世界黄金价格预测](#)
- [9. 基于BP神经网络的股票价格预测研究](#)
- [10. 基于RBF-BP神经网络的管道腐蚀类型预测](#)
- [11. 基于BP神经网络的居民消费价格预测研究](#)
- [12. 分段bezier曲线在黄金价格预测中的应用](#)
- [13. 基于BP神经网络模型的玉米价格基差预测](#)
- [14. 国际黄金价格的预测——基于VAR模型](#)
- [15. 基于ARFIMA-GARCH模型族的黄金价格预测分析](#)
- [16. BP网络和RBF网络在期货预测中的比较研究](#)
- [17. BP和RBF网络的结构动力响应预测对比](#)
- [18. 改进RBF神经网络在我国大豆价格预测中的应用研究](#)
- [19. 基于ARMA-GARCH模型的黄金现货价格预测](#)
- [20. 一种MPSO-BP-RBF网络模型及其在石油储层预测中的应用](#)
- [21. 基于BP神经网络的国际黄金价格预测模型](#)
- [22. 基于BP神经网络的黄金期货价格预测](#)
- [23. 2010年预测之黄金价格](#)
- [24. 基于BP网络和RBF网络的河道浅滩演变预测模型](#)
- [25. BP模型在强度预测中的应用](#)
- [26. 基于异质主体模型与RBF神经网络的黄金期货价格预测](#)
- [27. 基于BP神经网络的居民消费价格预测研究](#)
- [28. 基于Q-RBF神经网络模型的国产大豆价格预测研究](#)
- [29. 灰色理论与RBF神经网络模型在证券价格预测中的应用](#)
- [30. GA-BP算法在黄金价格预测中的应用](#)
- [31. 基于BP网络的房地产价格预测](#)

32. BP神经网络和RBF神经网络模型对图书流通量的预测分析
33. 基于PCA-BP神经网络的股票价格预测
34. BP神经网络模型与灰色GM(1,1)模型在需水预测中的应用
35. 创业板研发投入的神经网络预测:是采用RBF模型还是BP模型?
36. BP神经网络预测模型在高铁沉降预测中的应用
37. 改进BP网络模型在径流预测中的应用
38. 基于RBF-BP协作神经网络的公交行程时间预测
39. 基于BP人工神经网络的预测模型在基坑变形预测中的应用
40. 基于RBF-BP组合神经网络的地震预测研究
41. BP和RBF网络模型在黄金价格预测中的应用
42. 基于BP神经网络的黄金期货价格预测
43. BP神经网络的股票价格预测模型建立
44. 基于BP神经网络改进的黄金价格预测
45. 股票价格预测:GARCH模型与BP神经网络模型比较
46. 基于BP神经网络的黄金价格预测分析
47. 神经网络在股市预测中的应用——基于BP,RBF与GRNN
48. 基于PSO-BP与RBF神经网络的蔬菜价格组合预测
49. 基于差值灰色RBF网络模型股票指数预测
50. RBF神经网络模型及其在年径流预测中的应用