

兰州大学

硕士学位论文

非平稳时间序列建模与预测

姓名：管鹏

申请学位级别：硕士

专业：通信与信息系统

指导教师：吕振肃

20070601

摘 要

时间序列按统计特性可以分为平稳时间序列和非平稳时间序列两类。在实际生活中,我们经常遇到的序列,特别是反映社会、经济现象的序列,大多数并不平稳,而如果能够正确的预测这些序列,又可以对社会、经济的发展起到很好的控制和指导作用。因此研究非平稳时间序列的建模与预测具有很重要的现实意义。本文的重点是对非平稳时间序列的建模与预测方法进行研究。

在研究四种非平稳时间序列建模与预测方法的基础上,对其中的两种算法提出改进:一是在小波分析时对逼近系数的重构信号使用 *ARIMA* 法建模,使模型更符合信号的实际情况;二是对经过 *Kalman* 滤波后分解出趋势项的使用 *ARMA* 法建模,再利用模型外推预测从而提高了预测精度。另外对采用灰度预测模型和 *BP* 神经网络的非平稳时间序列预测方法也进行了研究。

四种算法都用 *Matlab* 编写了相应程序,以上证综合指数的预测为例,通过实验数据的比较,对四种算法进行了对比,总结了各自的优缺点,证明改进后的基于小波分析的非平稳时间序列的预测方法和 *Kalman-ARMA* 预测法在预测精度和曲线拟合度方面要优于灰度组合预测法和 *BP* 神经网络预测法。

关键词: 非平稳时间序列建模; 非平稳时间序列预测; 小波分析; *Kalman* 滤波; 灰度理论; *BP*神经网络

ABSTRACT

According to the characteristic of statistics, time series is divided into two classes. One is stationary time series, the other is non-stationary time series. In daily time we usually observe time series which is almost non-stationary especially in the phenomena of society and economic. Forecasting these series correctly can control and direct the advancement of society and economic greatly. So the research of the modeling and the forecasting of non-stationary time series is very important in practice. The study narrates mainly about the method of the modeling and the forecasting of non-stationary time series.

Based on the study of four methods of non-stationary time series modeling and forecasting, two algorithmic means are committed to be improved, one is using *ARIMA* modeling to reconstruction signals of approximation coefficients from a wavelet decomposition, it will more fit practice condition; two is using *ARMA* process to tendency vectors filtered from *Kalman* filter, it will improve forecasting precision. Furthermore, the study of using gray compounding model and *BP* neural network model is carried on.

Four methods are all actualized to software programs with *Matlab*. Based on experiment data of SSE(Shanghai Stock Exchange) Composite, four algorithmic means are compared each other, and summed up merits and disadvantages. It can be proved that the forecasting method based on wavelet decomposition and the method of *Kalman - ARMA* are better than gray compounding model and *BP* neural network model in forecasting precision and the fitting curve.

Keywords: non-stationary time series modeling; non-stationary time series forecasting; wavelet analysis; *Kalman* filter; gray theory; *BP* neural network

原创性声明

本人郑重声明：本人所呈交的学位论文，是在导师的指导下独立进行研究所取得的成果。学位论文中凡引用他人已经发表或未发表的成果、数据、观点等，均已明确注明出处。除文中已经注明引用的内容外，不包含任何其他个人或集体已经发表或撰写过的科研成果。对本文的研究成果做出重要贡献的个人和集体，均已在文中以明确方式标明。

本声明的法律责任由本人承担。

论文作者签名：管鹏 日期：2007.5.28

关于学位论文使用授权的声明

本人在导师指导下所完成的论文及相关的职务作品，知识产权归属兰州大学。本人完全了解兰州大学有关保存、使用学位论文的规定，同意学校保存或向国家有关部门或机构送交论文的纸质版和电子版，允许论文被查阅和借阅；本人授权兰州大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用任何复制手段保存和汇编本学位论文。本人离校后发表、使用学位论文或与该论文直接相关的学术论文或成果时，第一署名单位仍然为兰州大学。

保密论文在解密后应遵守此规定。

论文作者签名：管鹏 导师签名：吕振东 日期：2007.5.28

§ 1 绪 论

1.1 时间序列概述

按照时间的顺序把随机事件变化发展的过程记录下来就构成了一个时间序列。对时间序列进行观察、研究, 找寻它变化发展的规律, 预测它将来的走势就是时间序列分析。

从统计意义上讲, 所谓时间序列就是将某一个指标在不同时间上的不同数值, 按照时间的先后顺序排列而成的数列。

从数学意义上讲, 如果对某一过程中的某一个变量或一组变量 $X(t)$ 进行观察测量, 在一系列时刻 t_1, t_2, \dots, t_N ($i = 1, 2, \dots, N$) 为自变量, 且 $t_1 < t_2 < \dots < t_N$ 得到的离散有序数集合 $X(t_1), X(t_2), \dots, X(t_N)$ 称为离散数字时间序列, 即随机过程的一次样本实现。设 $X(t: t \in T)$ 是一个随机过程, $X(t_i) (i = 1, 2, \dots)$ 是在时刻 t_i 对过程 $X(t)$ 的观察值, 则 $X(t_i) (i = 1, 2, \dots)$ 称为一次样本实现, 也就是一个时间序列。

从系统意义上讲, 时间序列就是某一系统在不同时间(地点、条件等)的响应。这个定义从系统运行的观点出发, 指出时间序列是按一定顺序排列而成的; 这里的“一定顺序”既可以是时间顺序, 也可以是具有各种不同意义的物理量(如代表温度, 速度或其它单调递增取值的物理量)。可见, 时间序列只是强调顺序的重要性, 而并非强调必须以时间顺序排列。

时间序列根据所研究的依据不同, 有不同的分类。

1. 按所研究对象的多少来分, 有一元时间序列和多元时间序列。

2. 按时间的连续性可将时间序列分为离散时间序列和连续时间序列两种。

3. 按时间序列的统计特性来分, 有平稳时间序列和非平稳时间序列。如果一个时间序列的概率分布与时间 t 无关, 则称该时间序列为严格的(狭义的)平稳时间序列。如果时间序列的一、二阶矩存在, 而且对任意时刻 t 满足: (1) 均值为常数, (2) 协方差为时间间隔 τ 的函数, 则称该时间序列为宽平稳时间序列, 也叫广义平稳时间序列。反之, 把不具有平稳性的时间序列(即时间序列的均值或协方差是与时间有关的序列)称之为非平稳序列。

4. 按时间序列的分布规律来分, 有 Gaussian 型(Gaussian)时间序列和非 Gaussian 型(non-Gaussian)时间序列。

时间序列的一个重要的基本特征就是相邻观测值之间具有依赖性, 这种依赖特征具有很大的实际意义。时间序列分析所论及的就是对这种依赖性进行分析的技巧, 即根据已有的动态数据来揭示系统动态结构和规律。时间序列分析的基本思想是寻找系统的当前值与其过去的运行记录(观察数据)的关系(是一种纵向关系), 建立能够比较精确地反映时

间序列中动态依存关系的数学模型，并借此对系统的未来行为进行预报。

在实际问题中，我们常常遇到的序列，特别是反映社会、经济现象的序列，大多数并不平稳，而是具有明显的增长或减少趋势，或者含有依时间周期变化的趋势。例如国际航线客票数据、太阳黑子数据以及证券交易所每日收盘的综合指数等等。这就需要对非平稳时间序列进行研究。

1.2 时间序列预测

一般，从当前的和过去的观察值 $x(n), x(n-1), x(n-2), \dots$ 估计当前的信号值 $y(n) = \hat{s}(n)$ 称为过滤或滤波；从过去的观察值，估计当前的或将来的信号值 $y(n) = \hat{s}(n+N)$ ($N \geq 0$) 称为预测或外推；从过去的观察值，估计过去的信号值 $y(n) = \hat{s}(n-N)$ ($N \geq 1$) 称为平滑或内插。

时间序列预测方法的基本思想是：预测一个现象的未来变化时，用该现象的过去行为来预测未来。即通过时间序列的历史数据揭示现象随时间变化的规律，将这种规律延伸到未来，从而对该现象的未来做出预测。

时间序列一般是以上几种变化形式的叠加或组合。时间序列预测方法分为确定型和随机型两大类。确定型时间序列预测方法是用一个确定的时间函数 $y = f(t)$ 来拟合时间序列，不同的变化采取不同的函数形式来描述，不同变化的叠加采用不同的函数叠加来描述。随机型时间序列分析法是通过分析不同时刻变量的相关关系，揭示其相关结构，对时间序列进行预测。文中讨论的 *ARMA* 模型是一种确定型时间序列模型预测方法，其预测精度高于简单模型。在实际预测时，采用组合预测模型，可以达到更高的预测精度。

1.3 预测的评估标准

如果我们以 s 与 \hat{s} 分别表示信号的真值与估计值，而用 $e(n)$ 表示它们之间的误差，即

$$e(n) = s(n) - \hat{s}(n) \quad (1.1)$$

显然， $e(n)$ 可能是正的，也可能是负的，并且它是一个随机变量。因此，用它的均方值表达误差是合理的，所谓均方误差最小即它的平方的统计平均值最小：
即

$$E[e(n)^2] = E[(s(n) - \hat{s}(n))^2] = \frac{1}{N} \sum_{t=1}^N |x(t) - \hat{x}_N(t)|^2 \quad (1.2)$$

最小。

均方根误差是另一评估指标，所谓均方根误差就是均方误差的平方根。

$$S_N = \sqrt{E[e(n)^2]} = \sqrt{E[(s(n) - \hat{s}(n))^2]} \quad (1.3)$$

一个好的预测应该具有小的，独立的误差。在本文中都是以均方根误差作为评估标准。

采用均方根误差作为预测的评估标准的原因在于它的理论分析比较简单,不要求对概率的描述。并且在这种准则下导出的最佳线性系统对其他很广泛一类准则(如最小二乘法、极大似然估计和极大验后估计等)而言也是最优的。

1.4 时间序列建模、预测的研究现状

文献[6]中介绍的就是目前研究时间序列较有效的方法——Box-Jenkins建立的 *ARIMA* (Autoregressive Integrated Moving Average) 模型。它是以美国著名的统计学专家Box和英国的Jenkins的名字命名的一种时间序列预测方法。运用 *ARIMA* 模型的前提条件是,用作预测的时间序列是平稳序列,对于非平稳序列则可通过适当地变换(如差分、取对数)将它变为平稳序列,再利用 *ARIMA* 模型进行建模。但这种单一函数变换的方法将非平稳转化为平稳的能力是有限的,并且需要一定的技巧。

上世纪80年代后期小波理论发展并兴起后,由于小波变换可以将各种交织在一起的不同频率组成的混合信号分解成不相同频率的块信号,在时域和频域上具有良好的“显微镜”功能。原始信号通过小波分解一层一层地分解到不同的频率通道上,由于分解后的信号在频率成分上比原始信号单一,并且小波分解对信号做了平滑,因此分解后信号的平稳性比原始信号要好得多,对信息的分析研究具有明显的优势。文献[14]中介绍了一种基于小波分析的非平稳时间序列预测方法,过程是对原时间序列经过若干次小波分解,再对系数的重构序列采用传统的预测方法进行预测,各预测值相加得到原序列的预测值。这样就为非平稳时间序列预测提供了一种新的有效的方法。

文献[18][19]中介绍的近年兴起的一类很重要的非线性模型——BP神经网络模型。其基本思想是通过在解释变量和被解释变量之间插入神经元(或节点)来模拟大脑结构。对任何神经元输入的线性函数都有可能通过适当的函数(可以是非线性的)进行改造,从而给每个神经元赋值,然后对这些数值进行组合来近似地得到被解释变量值。在时间序列分析中,被解释变量就是需要预测的变量,解释变量就是适当的滞后变量。模型结构中通常隐含了许多参数,通过神经网络可以以黑箱方式进行比较好的模拟和预测。

文献[23][24]则介绍了一种由邓聚龙教授建立的灰色预测模型,这种模型具有所需原始数据不多,容易采集,方法简便并具有较高的准确性,在近二十年来也得到了广泛的应用。

在算法理论上取得进展的同时,非平稳时间序列预测的应用领域也在不断扩展。从国民经济、天气预报到市场分析、股指预测、故障诊断等领域中都有了广泛的应用。总的来说,所面临的形势是既令人兴奋,又富有挑战性。一方面,可供选择的模型越来越多,而且还有比较好的软件可供使用;另一方面,一些预测方法还处在不断发展和完善当中,较新的一些算法模型仍是探索性的,尚需在实践中接受检验和继续发展,因而离广泛的应用

还有很长的一段路要走，还需要研究人员做更多的努力。

1.5 目前研究存在的问题

目前，非平稳时间序列预测研究存在的主要问题有以下几个方面：

1) 算法的多样性，反而使选择算法变的更加困难，更容易选择错误的或不合适的模型，使预测不准确甚至无效。而选择何种模型往往需要大量的验算或更好的专业知识。

2) 虽然现代软件很多，但一般操作都比较复杂，制约了更多用户的使用，还有可能使用错误的预测软件，造成错误的判断。

3) 近来一些新兴的应用领域对预测精度提出了更高的要求，这就需要对一些传统算法进一步改进或发展出新的算法模型。

4) 很多算法是根据样本内拟合方法得到的，用此方法通常能得到很好的拟合图像，很高的预测精度。但样本外推预测值通常没有预期的那么好，比较不同预测方法的精度，还是应该以样本外推预测结果为依据。

5) 许多观测到的时间序列表现出非线性特征，尤其是长期金融数据时间序列，故非线性模型近来格外引人注目。不过根据非线性模型得到的样本外推预测结果并不总像人们所预期的那样有多大改进，而人们对神经网络模型的优点则有些夸大其词^[25]。

1.6 本文的研究工作和章节安排

结合当前国内外非平稳时间序列预测领域的研究现状，本论文以非平稳时间序列建模与预测为研究对象，结合上证综合指数，使用 *Matlab* 实现了四种非平稳时间序列预测的算法：第一种算法利用小波分解对非平稳时间序列进行平滑、分解，对分解系数的单支重构序列再采用经典的 *ARIMA* 模型建模，最后利用模型进行外推预测；第二种算法通过 *Kalman* 滤波过程分解出非平稳时间序列的状态变量，通过状态变量得到趋势项，对趋势项建立 *ARMA* 模型，再利用模型进行预测；第三种算法采用灰度预测模型提取非平稳时间序列中的趋势项，再对残差部分使用 *ARMA(p, q)* 模型建模，利用 *ARMA* 模型进行预测；第四种算法利用较新的BP神经网络直接对非平稳时间序列进行非线性建模，利用训练好的网络进行预测。文章的最后对四种算法进行了分析、比较，对比了各自的优缺点。

本文共分7章。第1章介绍了非平稳时间序列预测的研究背景和意义，并简述了非平稳时间序列预测的研究现状及存在的问题；第2章介绍了经典的 *ARIMA* 模型法建模与预测；第3、4、5、6章分别研究了上述的四种算法；最后通过第7章对这四种算法进行了比较和总结，并对未来的研究方向进行了展望。

§2 ARIMA 模型法

2.1 概述

ARIMA 模型又称 Box-Jenkins 模型, 是由美国统计学家 G. E. P. Box 和英国统计学家 G. M. Jenkins 于 1970 年首次系统提出的 (参见文献[4])。Box-Jenkins 方法是一种理论较为完善的统计预测方法, 他们的工作为实际工作者提供了对时间序列进行分析、预测, 以及对 *ARMA* 模型识别、估计和诊断的系统方法。使 *ARMA* 模型的建立有了一套完整、正规、结构化的建模方法, 并且具有统计上的完善性和牢固的理论基础。其优点在于一旦建立了精确的信号模型, 并确定模型的系数之后, 就可以根据有限的数据集对信号的趋势进行预测, 就好比获得了无限的采样数据一样, 因此可以大幅度地提高预测精度。但 *ARMA* 模型建立的前提是时间序列为平稳序列, 当时间序列为非平稳序列时, 文献[4]中提出可通过适当地变换 (如差分、取对数) 将它变为平稳序列后, 再利用 *ARMA* 模型进行建模, 这类模型 Box-Jenkins 称为求和自回归滑动平均 (*ARIMA*) 模型。

2.2 平稳时间序列模型

2.2.1 AR、MA、ARMA 模型

序列 $x(n)$ 的 p 个极点 q 个零点的 $ARMA(p, q)$ 模型可以描述为:

$$x(n) = -\sum_{k=1}^p a_k x(n-k) + \sum_{k=0}^q b_k u(n-k) \quad (2.1)$$

并假定 $b_0 = 1$ 。

对 (2.1) 式可分三种情况讨论:

1、如果 b_1, b_2, \dots, b_q 全为零, 则 (2.1) 式变为

$$x(n) = -\sum_{k=1}^p a_k x(n-k) + u(n) \quad (2.2)$$

此模型称为自回归 (Auto-regressive) 模型, 简称 AR 模型, 该模型现在的输出是现在的输

入和过去 p 个输出的加权和。

2、如果 a_1, a_2, \dots, a_p 全为零, 则(2.1)式变为

$$x(n) = \sum_{k=0}^q b_k u(n-k) = u(n) + \sum_{k=1}^q b_k u(n-k), \quad b_0 = 1 \quad (2.3)$$

此模型称为滑动平均 (Moving-Average) 模型, 简称 MA 模型。

3、如果 $a_1, a_2, \dots, a_p; b_1, b_2, \dots, b_q$ 不全为零, 则(2.1)式给出的模型称为自回归—滑动平均 (Auto-regressive Moving-Average) 混合模型, 简称 ARMA 模型。

AR, MA 和 ARMA 是平稳时间序列最主要的参数模型。AR 模型的正则方程是一组线性方程, 而 MA 和 ARMA 模型是非线性方程。Wold 分解定理告诉我们任何有限方差的 ARMA 或 MA 平移过程可以用可能是无限阶的 AR 模型表达; 同样, 任何 ARMA 或 AR 模型可以用可能是无限阶的 MA 模型表达。因此, 如果在这三个模型中选了一个与信号不匹配的模型, 但只要模型的阶足够高, 它任然能够比较好地逼近被建模的随机过程^[2]。三种模型中 AR 模型具有一系列好的性能, 因此, 是被研究最多并获得广泛应用的一种模型^[1]。

2.2.2 时间序列的自相关和偏自相关函数

对于模型预测, 首先要找到与其拟合最好的预测模型, 其次模型阶数的确定和参数的估计也是建模的关键。确定模型类型最重要的工具是序列的自相关和偏自相关函数 (参见文献[5], [11], [13])。

1、自相关函数 (ACF)

构成时间序列的每个序列值 $x_t, x_{t-1}, \dots, x_{t-k}$ 之间的简单相关关系称为自相关。自相关程度由自相关系数 r_k 度量, 表示时间序列中相隔 k 期的观测值之间的相关程度。

$$r_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.4)$$

其中, n 是样本量, k 为滞后期, \bar{x} 代表样本数据的算术平均值。

与简单相关系数一样, 自相关系数 r_k 的取值范围是 $[-1, 1]$, 并且 $|r_k|$ 越接近 1, 自相关程度越高。

2、偏自相关函数 (PACF)

偏自相关是指对于时间序列 x_t , 在给定 $x_{t-1}, x_{t-2}, \dots, x_{t-k+1}$ 的条件下, x_t 与 x_{t-k} 之间的条件相关关系。其相关程度用偏自相关系数 φ_{kk} 度量, 有 $-1 \leq \varphi_{kk} \leq 1$ 。

$$\varphi_{kk} = \begin{cases} r_1 & k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} \varphi_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \varphi_{k-1,j} r_j} & k = 2, 3, \dots \end{cases} \quad (2.5)$$

在实际应用中,应该综合考察序列的自相关与偏自相关。将时间序列的自(偏自)相关系数绘制成图,并标出一定的随机区间,称为自(偏自)相关分析图。它是对时间序列进行自(偏自)相关分析的主要工具。

2.2.3 模型类型的判别

1、AR、MA、ARMA三种模型自相关与偏自相关函数的特点

(1) $AR(p)$ 模型

$AR(p)$ 模型的偏自相关函数为:

$$\varphi_{kj} = \begin{cases} \varphi_j & 1 \leq j \leq p \\ 0 & p+1 \leq j \leq k \end{cases} \quad (2.6)$$

也就是说, $AR(p)$ 模型的偏自相关函数 φ_{kj} 是 p 步截尾的, 即当 $j > p$ 时, φ_{kj} 的值为0。这个特点对识别自回归模型以及确定阶数 p 都很有用。而 $AR(p)$ 模型的自相关函数呈指数或者正弦波衰减, 具有拖尾性。

(2) $MA(q)$ 模型

$MA(q)$ 模型的自相关函数为

$$r_k = \begin{cases} \sigma^2 \sum_{k=m}^q b(k)b(k-m) = \frac{-b_k + b_1 b_{k+1} + \dots + b_{q-k} b_q}{1 + b_1^2 + b_2^2 + \dots + b_q^2} & 1 \leq k \leq q \\ 0 & k > q \end{cases} \quad (2.7)$$

则 $MA(q)$ 模型的自相关函数 r_k 在 $k > q$ 以后全部是0, 即自相关函数是 q 步截尾的。而与 $AR(p)$ 模型相反, $MA(q)$ 模型的偏自相关函数随着滞后期 k 的增加, 呈现指数或者正弦波衰减, 趋向于0, 是拖尾的。

(3) $ARMA(p,q)$ 模型

$ARMA(p,q)$ 模型的自相关函数和偏自相关函数均是拖尾的^[11,21,22]。

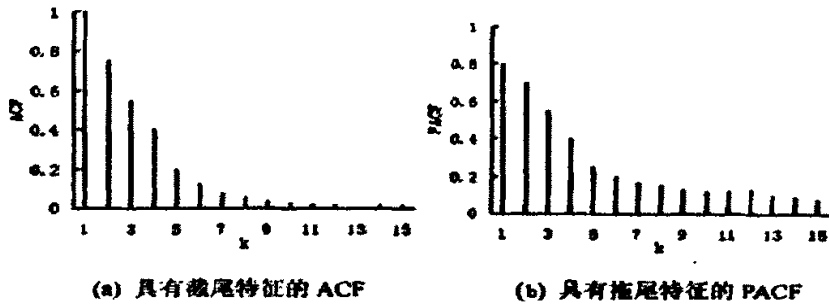


图2.1 截尾ACF与拖尾PACF图

2、由自相关与偏自相关函数确定模型类型

根据以上 $AR(p)$ 、 $MA(q)$ 、 $ARMA(p,q)$ 序列的自相关和偏自相关函数的统计特性，可以初步确定模型的类型：

表2.1 通过自相关函数与偏自相关函数确定模型类型

模 型	$AR(p)$	$MA(q)$	$ARMA(p,q)$	$ARIMA(p,d,q)$
自相关函数	拖尾	截尾	拖尾	拖尾（呈现非指数衰减）
偏自相关函数	截尾	拖尾	拖尾	拖尾（呈现非指数衰减）

2.2.4 模型阶数的选择

1、基于自相关函数和偏相关函数的定阶方法

对于 $ARMA(p,q)$ 模型，可以利用其样本的自相关函数 $\{r_k\}$ 和样本偏自相关函数 $\{\varphi_{kk}\}$ 的截尾性判定模型的阶数（参见文献[11]，[21]，[22]）。

具体方法如下：

(1) AR模型，

当 $k > p$ 时， $\varphi_{kk} = 0$ ，对于 $k > p$ ， N 充分大， φ_{kk} 为渐进正态 $N\left(0, \frac{1}{N}\right)$ ，故

$$P\left(|\varphi_{kk}| \leq \frac{1}{\sqrt{N}}\right) = 68.3\%, \text{ 或 } P\left(|\varphi_{kk}| \leq \frac{2}{\sqrt{N}}\right) = 95.5\%$$

若对于某正整数 n_0 ， $\varphi_{n_0 n_0}$ 明显地异于零，而 φ_{n_0+1, n_0+1} ， φ_{n_0+2, n_0+2} ， \dots ， φ_{n_0+M, n_0+M} 均近似于零，且满足 $|\varphi_{kk}| \leq \frac{1}{\sqrt{N}}$ （或 $|\varphi_{kk}| \leq \frac{2}{\sqrt{N}}$ ）的个数占 M 的比例近似的为 68.3%（或 95.5%），则可以近似的判定 $\{\varphi_{kk}\}$ 是在 n_0 步截尾，因而初步判定 AR 模型的阶数为 n_0 ，即认为平稳时间序列 $\{x_t\}$ 为 $AR(p_0)$ 模型。

(2) MA模型

类似,我们可通过计算序列 $\{r_k\}$,考察其中满足 $|r_k| \leq \frac{1}{\sqrt{n}}$ 或者 $|r_k| \leq \frac{2}{\sqrt{n}}$ 的个数是否占M个的68.3%或者95.5%。即可以近似的判定 $\{r_k\}$ 是 q_0 步截尾,平稳时间序列 $\{x_t\}$ 为 $MA(q_0)$ 。

(3) ARMA 模型

如果对于序列 $\{\varphi_{ik}\}$ 和 $\{r_k\}$ 来说,均不截尾,即不存在上述的 p_0 和 q_0 ,则可以判定平稳时间序列 $\{y_t\}$ 为ARMA模型。

2、利用信息准则法定阶

文献[1]提出了几种不同的信息准则确定模型阶次:

(1) 最终预测误差(FPE)准则:

$$FPE(k) = \hat{\rho}_k \frac{N + (k + 1)}{N - (k + 1)} \quad (2.8)$$

(2) Akaike 信息准则(AIC)

$$AIC(k) = N \ln(\hat{\rho}_k) + 2k \quad (2.9)$$

(3) BIC准则函数:

$$BIC(k) = N \ln(\hat{\rho}_k)^2 + k \ln N \quad (2.10)$$

式中 N 为数据 $x_N(n)$ 的长度,当阶次 k 由1增加时,上述准则都将在某一个 k 处取得极小值,将此时的 k 定为最合适的阶次。

上述三式仅为阶次的选择提供了一个依据,在实际运用中,基于自相关函数和偏相关函数的定阶方法更为准确。

2.2.5 模型参数的求解

1、AR 模型

在利用AR模型进行预测时,必须计算出AR模型的参数 a_1, a_2, \dots, a_p 及白噪声序列的方差 σ^2 。这些参数的提取算法主要包括自相关法、Burg算法、协方差、改进的协方差法,以及最大似然估计方法等(参见文献[1])。

上述方法中,改进的协方差的估计性能最好,但计算比较复杂;自相关法的计算最为简单,但估计的分辨率相对较差;Burg方法较为通用,计算不太复杂,且具有较好的估计质量[3]。

2、MA 模型

MA模型参数的求解方法有三种:(1)、谱分解法;(2)、Durbin法(用高阶的AR模型

来近似 MA 模型); (3)、最大似然估计。其中最有效的求解方法是 Durbin 法^[1]。

3、ARMA模型

文献[1]介绍了ARMA模型参数的求解的三个步骤:

(1) 利用AR模型参数算法估计AR参数 $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_p$;

(2) 对已知数据 $x_N(n)$, 用FIR滤波器 $\hat{A}(z) = 1 + \sum_{k=1}^p \hat{a}_k z^{-k}$ 滤波, 那么滤波器的输出 $y(n)$ 将近似一个 $MA(q)$ 过程:

(3) 用Durbin法求解 $MA(q)$ 参数, 求出 $\hat{b}_1, \hat{b}_2, \dots, \hat{b}_q$, 从而实现 $ARMA(p, q)$ 模型的参数估计。

2.2.6 平稳时间序列预测

1、AR(p)模型预测^[1]

利用 n 之前的 p 个值对 $x(n)$ 作预测, 称之为“前向预测”, 记为:

$$\hat{x}_t^f(m) = \begin{cases} -\sum_{i=1}^p a^f(i)x(t+l-i) & m=1 \\ -\sum_{i=1}^{m-1} a^f(i)\hat{x}_t^f(m-i) - \sum_{i=m}^p a^f(i)x(t+m-i) & 1 \leq m \leq p \\ -\sum_{i=1}^p a^f(i)\hat{x}_t^f(m-i) & m > p \end{cases} \quad (2.11)$$

上标 f 表示前向预测(forward prediction)。 $\hat{x}_t^f(m)$ 表示在 t 时刻 m 步的前向预测。利用自相关法、Burg 算法、协方差、改进的协方差法等方法得到模型的参数后, 就可以进行前向预测, 利用预测值递推可依次得到多步预测值。

2、MA(q)、ARMA(p,q)模型预测

$MA(q)$ 、 $ARMA(p, q)$ 的外推预测一般都是将 $MA(q)$ 、 $ARMA(p, q)$ 模型转换为相应的高阶 AR 模型, 再利用 AR 模型的预测公式进行外推预测。

3、预测误差

预测误差公式为:

$$e_t(l) = x_{t+l} - \hat{x}_t(l) = \psi_0 \varepsilon_{t+l} + \psi_1 \varepsilon_{t+l-1} + \dots + \psi_{l-1} \varepsilon_{t+1} \quad (2.12)$$

l 步线性最小方差预测的方差和预测步长 l 有关, 而与预测的时间原点 t 无关。预测步长 l 越大, 预测误差的方差也越大, 因而预测的准确度就会降低。所以一般不能用 AR、MA 和 ARMA 模型作为长期预测模型。

2.3 非平稳时间序列模型

ARIMA 模型意为求和自回归滑动平均模型 (Integrated Auto-regressive Moving Average Model), 简记为 *ARIMA*(p, d, q)^[4], p, q 分别为自回归和滑动平均部分的阶次, d 为差分运算阶次, 对于某些非平稳时间序列 $\{y(t)\}$, 其一般形式为:

$$\varphi(B)(1-B)^d y(t) = \theta(B)a(t) \quad (2.13)$$

若将 $(1-B)^d y(t)$ 记为 $z(t)$, 则上式就是 *ARMA* 模型。

Box发现, 可通过差分方法求出增量序列: $\nabla y(t) = y(t) - y(t-1)$ ($t = 2, 3, \dots, N$)。经过一次差分后, 如果此增量序列 $\{\nabla y(t)\}$ 是平稳的, 那么对 $\{\nabla y(t)\}$ 建立 *ARMA* 模型, 表示为:

$$\varphi(B)\nabla y(t) = \theta(B)a(t) \quad (2.14)$$

根据差分算子 ∇ 与后移算子 B 的关系 ($\nabla = 1 - B$), 即 $(1-B)y_t = y_t - y_{t-1}$, 有

$$\varphi(B)(1-B)y(t) = \theta(B)a(t) \quad (2.15)$$

即还原得到 $\{y(t)\}$ 的 *ARIMA* 模型。

以上对非平稳时间序列 $\{y(t)\}$ 作一次差分称为一阶差分。将这种思路推广, 当采用一阶差分还不能使 $\{\nabla y(t)\}$ 成为平稳时间序列时, 还可采用高阶 (d 阶) 差分, 以使 $\{\nabla^d y(t)\}$ 成为平稳时间序列, 再对 $\{\nabla^d y(t)\}$ 建立 *ARMA* 模型, 然后根据差分算子 ∇ 与后移算子 B 的关系 ($\nabla = 1 - B$), 得到非平稳时间序列 $\{y(t)\}$ 的 *ARIMA* 模型, 这就是 *ARIMA* 模型法的基本思路。虽然足够多次的差分运算可以充分地提取原序列中的非平稳确定性信息, 但过度的差分也会造成有用信息的浪费。一般而言, 若某时间序列具有线性的趋势, 则可以对其进行一次差分而将线性趋势剔除掉; 若某时间序列具有指数的趋势, 则可以取对数将指数趋势化为线性趋势, 然后再进行差分以消除线性趋势, 接着对差分后的序列拟合 *ARMA* 模型进行分析与预测, 最后再通过差分的反运算得到 $\{y(t)\}$ 预测值。

ARIMA 模型法的流程图如图 2.2 所示:

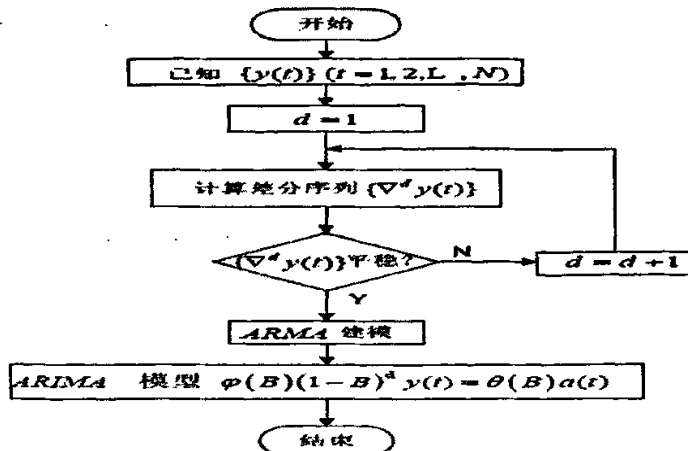


图 2.2 *ARIMA* 模型法流程图

§3 基于小波分解的非平稳时间序列建模与预测

3.1 概述

小波变换是一种信号的时间—尺度分析方法，它不仅继承和发展了短时傅立叶变换的局部化的思想，而且克服了窗口大小不随频率变化，缺乏离散正交基的缺点，是一种比较理想的信号处理的方法。小波变换具有多分辨率分析的特点，而且在时、频两域都具有表征信号局部特征的能力，是一种窗口大小固定不变但形状可改变，时间窗和频率窗都可以改变的时频局部化分析方法。即在低频部分具有较高的频率分辨率，在高频部分具有较高的时间分辨率和较低的频率分辨率，很适合于探测正常信号中夹带的瞬态反常现象并展示其成分，所以被誉为分析信号的显微镜。

时间序列 $x(t)$ ，一般由趋势项 $d(t)$ ，周期项 $s(t)$ 和随机干扰项 $y(t)$ 组成，

$$x(t) = d(t) + s(t) + y(t) \quad (3.1)$$

其中趋势项反映的是时间序列的主要特征，例如在经济领域中，趋势项代表的是经济变量季节性趋势或多年变化趋势，周期项则反映序列的周期性变化。把趋势项分离出来之后，周期项 $s(t)$ 和剩下的随机项 $y(t)$ 可以用均值为零的平稳时间序列来分析。

由小波分析理论可知，信号可以通过小波分解一层一层分解到不同的频率通道上，表现为由于分解后的信号在频率成分上比原始信号单一，细节项常为周期振荡信号，可以用均值为零的平稳时间序列来分析；逼近系数则表示趋势或走向，已被小波分解对作了平滑，因此平稳性比原始信号好得多，一般小波算法此时就将逼近系数作为平稳序列进行处理（参见文献[14]）。但多数情况下逼近系数此时依然为非平稳信号，因此我在此处作了改进，采用 $ARIMA(p, d, q)$ 法建模，利用差分方程使逼近系数平稳后再建立 $ARMA$ 模型。这样就将小波分析与经典的平稳时间序列建模结合起来，从而为非平稳时间序列的建模与预测提供了一种新的方法，并且提高了预测精度。

3.2 基于小波分解的非平稳时间序列建模与预测

3.2.1 小波分解

对非平稳时间序列 $\{x_t\}$ ($t = 1, 2, \dots, N$)，在利用 $ARIMA(p, d, q)$ 模型建模之前，需先对 $\{x_t\}$ 进行小波分解。

1、分解方式的选取

利用小波对时间序列进行分解与合成时，有两种方法可以采用。一是用二进小波变换

及其逆变换,这种方法分解得到的两个序列均与原时间序列长度相同,但不存在分解尺度的选择。其二是基于Mallat算法的多尺度小波分解与合成(参见文献[6]),这种方法每分解一次得到的两个序列均是分解前时间序列长度的一半,通过单支重构可使各分解序列恢复到原序列长度。这种方法可以得到不同分解尺度下的细节系数和逼近系数,相当于将信号分解成了各特定频率范围内的时间信号,因而适合对信号的分析、建模和预测。本文采用这种方法。

2、基小波的选取

小波基函数的选择是小波分解的关键。目前,已存在好几种不同的基小波,如Haar小波、Daubechies(dbN)小波系、Meyer小波和Symlets小波等。由于不同的小波基在正交性、紧支性、平滑性甚至对称性上表现出不同的特性,对于同一信号,采用不同的小波基函数的分析效果是不同的。

实际预测过程中,可根据不同的问题选择不同的小波母函数,同时结合考虑不同小波母函数的不同特性对预测值的影响,通过分析比较各小波函数处理信号的结果并与理论结果相比较,用误差来判定所选的小波函数的好坏。表3.1列出了决定小波函数质量的主要性质及一些常用小波函数性质^[9]:

表3.1 常用小波函数的一些性质

小波函数	Haar	Daubechies	Symlets	Meyer
正交性	有	有	有	有
紧支撑性	有	有	有	无
支撑长度	1	2N-1	2N-1	有限长度
对称性	对称	近似对称	近似对称	对称

这些性质的影响如下:

(1) 正交性。它对信号的重构获得较好的平滑效果是非常重要的。

(2) 紧支撑性。它决定了函数再经过小波分解后,反映了对称性及正交性两方面的继承性。

(3) 支撑长度。它决定了函数的局部优良性,以及频率分辨率的高低。而且进一步影响对称性的特性。

(4) 对称性。它能降低量化的误差,能保证小波的滤波特性有线性相移,不会造成信号失真,对于避免移相非常有用。

Daubechies已证明,具有紧支撑和对称性的正交小波仅有Haar小波,但Haar小波是不连续的,频域局部性差,结构简单,常应用于理论研究中,而且Haar小波是 $p=1$ 的Daubechies小波^[7]。而根据Daubechies小波的正交、时频紧支撑、高正规性和具有Mallat快速算法等特点,对于非平稳序列的分解具有很好的特性,因此本文采用Daubechies小波。

3、小波分解层数的确定

小波分解的特性是越往下分解,信号的频率段划分得越细,细节信号和逼近信号的平滑性和平稳性也越好。但是由于在分解过程中必然会存在计算上的误差,分解层数越多,

误差越大, 计算上的误差会带到预测的误差中, 使预测精度下降。因此在选择分解层数时, 不宜过少, 也不宜过多。实践中主要通过用小波分析方法处理信号的结果与理论结果的误差来判定分解层数。

4、多尺度小波分解

确定基小波和分解层数后, 利用Mallat算法将 $\{x_t\}$ 分解至 J 层。如下:

$$\begin{cases} c_{j+1} = Hc_j \\ d_{j+1} = Gc_j \end{cases}, \quad j = 0, 1, \dots, J-1 \quad (3.2)$$

式中, H 和 G 分别为一低通滤波器和一高通滤波器, J 为最大分解层数, d_1, d_2, \dots, d_J 为各层的细节系数, $c_0, c_1, c_2, \dots, c_J$ 为逼近系数, c_0 时信号尚未分解, 所以 $c_0 = \{x_t\}$ 。小波分解的过程如下图所示:

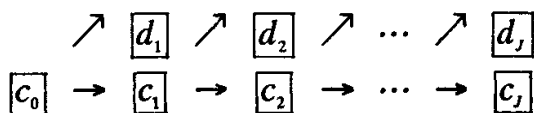


图3.1 多尺度小波分解图

通过小波分解将原始序列 $\{x_t\}$ 分解为细节系数 d_1, d_2, \dots, d_J 和逼近系数 c_J 。

3.2.2 小波系数的单支重构

采用Mallat算法进行小波分解后, 每一次分解后得到的细节系数和逼近系数比分解前的序列点数减少一倍, 点数的减少对预测是不利的。但分解系数可以通过单支重构恢复到原序列长度, 同时对每个小波分解系数的单支重构也可以提取各频带范围的时域信号^[9]。

设总信号为 $\{x_t\}$, 则

$$\{x_t\} = D_1 + D_2 + \dots + D_J + C_J = \sum_{j=1}^J D_j + C_J \quad (3.3)$$

式中, $D_1, D_2, \dots, D_J, C_J$ 为各层分解系数的单支重构信号。

单支重构的实质就是保持尺度系数一定, 仅对平移因子进行积分, 得到在特定尺度下的单支时间信号。单支重构的意义在于将信号分解成各特定频率范围内的时间信号, 对这种单支重构信号进行分析, 可以排除其他频率成分的干扰。

对 $D_1, D_2, \dots, D_J, C_J$ 分别建模、预测, 相加后就可以得到总信号 $\{x_t\}$ 的预测值。

3.2.3 建模及预测

1、细节系数重构信号的建模及预测

重构后的细节系数一般可视为平稳过程来处理, 可用 $AR(p)$ 、 $MA(q)$ 或 $ARMA(p, q)$ 等

模型进行预测。这里我们以AR模型来说明这种预测。

(1)、对每个 $D_s: \{d_{s1}, d_{s2}, \dots, d_{sM}\}, 1 \leq s \leq J$ 分别建立AR模型, 共计 J 个AR模型, 即令

$$d_{s,t} = \varphi_1 d_{s,t-1} + \varphi_2 d_{s,t-2} + \dots + \varphi_M d_{s,t-M} + \alpha_t, \quad \{t=1, 2, \dots, M\}, \{1 \leq s \leq J\} \quad (3.4)$$

并用已知的 $x_i (i \leq M)$ 分别对这 J 个AR模型进行参数估计和模型检验。

(2)、对 $AR(p)$ 模型 $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \alpha_t$ 而言, 其前向预测公式为:

$$\hat{x}(l) = \begin{cases} -\sum_{i=1}^p \varphi_i x_{t+l-i} & l=1 \\ -\sum_{i=1}^{l-1} \varphi_i x_{t+l-i} - \sum_{i=l}^p \varphi_i \hat{x}_{t+l-i} & 1 < l \leq n \\ -\sum_{i=1}^p \varphi_i \hat{x}_{t+l-i} & l > n \end{cases} \quad (3.5)$$

在上述经过模型检验的 J 个AR模型中使用这个最佳预测公式, 分别对每个

$d_{s,M+k} (s=1, 2, \dots, J)$ 进行预测, 得到预测值 $\hat{d}_{s,M+k} (s=1, 2, \dots, J)$ 。

2、逼近系数重构信号的建模及预测

逼近系数表示序列的趋势或走向, 一般为非平稳序列, 因此不能直接采用 $ARMA$ 模型, 需要经过差分令其平稳后, 方可采用 $ARMA$ 建模。即逼近系数采用的应该是 $ARIMA(p, d, q)$ 建模法。

通过 $ARIMA(p, d, q)$ 建模后, 逼近系数重构信号便可以采取与上述相同的预测方法, 得到 C_J 的预测值 $\hat{C}_{J,M+k}$ 。则原序列 x_t 在 M 时刻 k 步的预测值为:

$$\hat{x}_{M+k} = \hat{d}_{1,M+k} + \hat{d}_{2,M+k} + \dots + \hat{d}_{J,M+k} + \hat{C}_{J,M+k} \quad (3.6)$$

这就是我们想要得到的预测值。当然, 计算数学的一个新算法的产生还必须和其它的已有的算法进行比较, 才能真正体现自己的价值。下面我们以上证综合指数为样本进行预测与观察。

3.3 在上证综合指数预测中的应用

金融数据从本质上讲也是一种时间序列, 和我们平常分析的信号具有相同的特性。一般来说, 信号是指包含信息的物理量, 且这些物理量都是能够观测的。金融数据同样可看成是包含若干信息的物理量, 且能由观测得到的。金融时间序列作为一种特殊的时间序列, 通常具有如下三大特点: (1) 产生过程的随机、复杂性; (2) 多数含有噪声; (3) 数据间具有较强的非线性。

为了在对上证综合指数进行预测的同时对基于小波分解的非平稳时间序列建模方法的预测效果进行评价和比较，我们将整个样本数据分成估计样本与评价样本两部分。估计样本(从2006.1.4 到2006.12.29 共229个交易日)用于对模型的参数估值，而评价样本(2007.1.4 到2007.1.31 共20个交易日)用来对模型的预测结果进行检验。

图3.1所示是估计样本229个交易日的股价走势图，数据来源为“雅虎财经”。

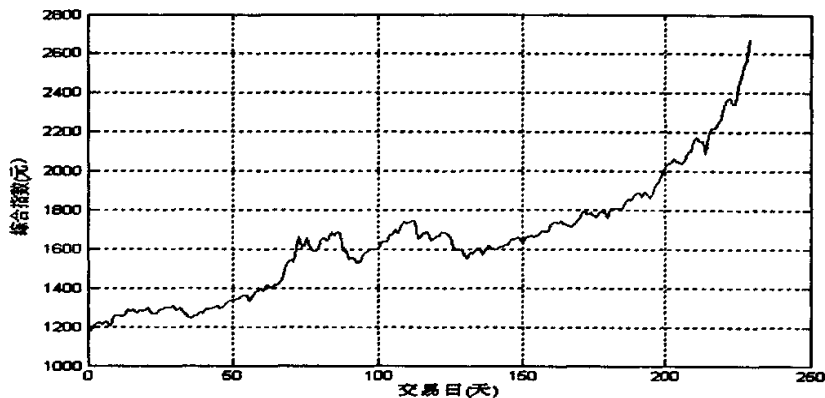


图 3.1 上证综合指数 2006 年走势图

3.3.1 数据的小波分解

1、选取基小波

根据3.2.1中介绍，决定采用Daubechies小波系作为小波基函数。但不同N值的Daubechies小波，可以得到不同的处理效果。具体采用何种N值则需要通过用小波分析方法处理信号的结果与实际结果的误差均方根值来判断。

表3.2给出了选择Daubechies小波系中db1~db10时采用小波分析预测方法得到的各预测误差均方根值。

表3.2 选取不同小波基函数时的预测结果比较

小波分解层数	db1	db2	db3	db4	db5	db6	db7	db8	db9	db10
预测误差均方根值	244.6204	248.0244	151.0825	209.9196	95.6065	173.3004	172.6966	189.0716	260.0721	243.9707

由上表可知， $N = 5$ 时得到的预测误差均方根值最小，因此采用 $N = 5$ 的Daubechies小波作为基小波。

2、确定小波分解层数

同样，分解层数的选择也需要通过比较误差均方根值来判断。

表3.3给出了采用Daubechies小波在分解层数为1~9时得到的各预测误差均方根值。

表3.3 不同分解层数时的预测结果比较

小波分解层数	1	2	3	4	5	6	7	8	9
预测误差均方根值	232.0087	204.0432	99.8059	98.2436	95.6065	296.1078	260.2087	264.1951	306.1681

由上表可知,当分解层数为3~5时,预测的误差基本达到了最小,因此在选择分解层数时,如果待预测的时间序列数据量不是很大,则分解层数一般选择3~5。本文选择分解层数为5。

3、小波分解

确定基小波和分解层数后,就可以将原始序列 $\{x_i\}$ 分解为细节系数 d_1, d_2, \dots, d_J 和逼近系数 c_J 。

3.3.2 综合指数小波分解系数的单支重构

对细节系数 d_1, d_2, \dots, d_J 和逼近系数 c_J 分别进行单支重构,得到单支重构信号 D_1, D_2, \dots, D_J 和 C_J ,如图3.2所示。

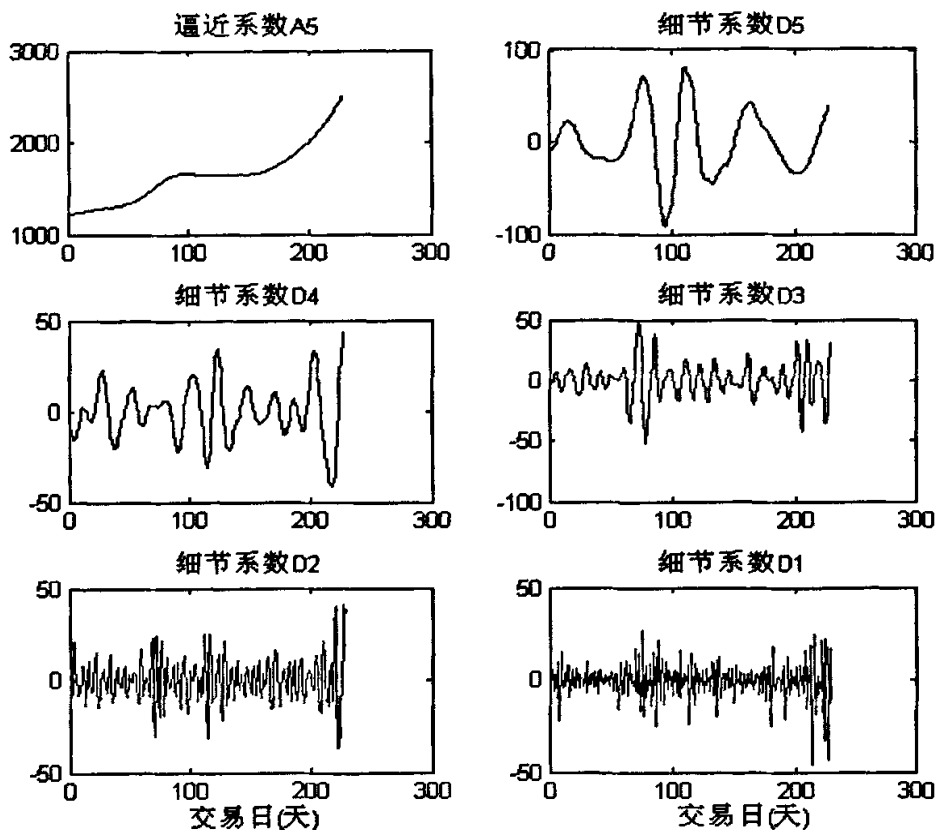


图3.2 小波分解各系数单支重构走势图

如图所示,相对原序列逼近系数已平滑许多。

3.3.3 逼近系数建模

通过比较可以看出进行小波分解后,细节系统基本上为均值为零的周期振荡信号,可以用平稳时间序列模型来分析和预测。而逼近系数仍为非平稳时间序列(如图3.3所示),

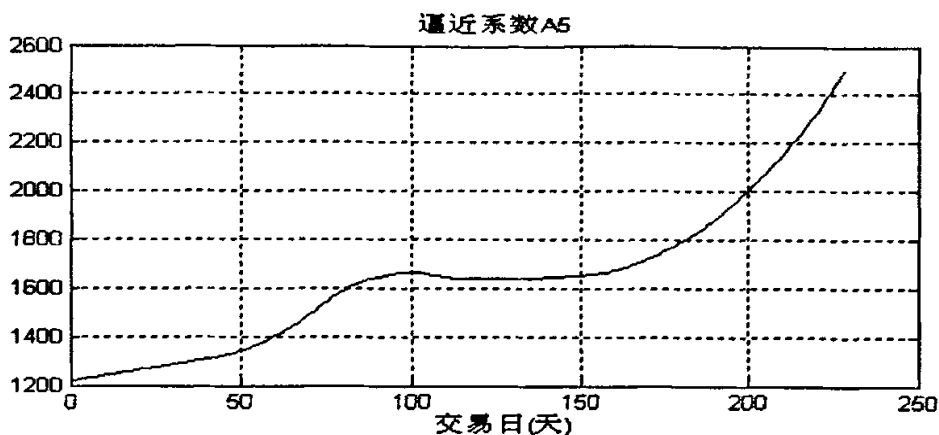


图 3.3 逼近系数单支重构走势图

需要使用差分方程令其平稳后再建模。即对原始的序列 $\{x_t\}$ 进行 d 次差分（一般不超过两次），得到一个平稳时间序列 $\{y_t\}$ ，再用 y_t 减去序列 $\{y_t\}$ 的均值 \bar{y} 得到均值为0的平稳序列 $\{y'_t\}$ 。然后采用Box-Jenkins方法建立相应的模型。

1、确定模型类别

图3.4为使用EView软件分析逼近系数的自相关和偏相关函数：

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.973	0.973	219.71	0.000
		2	0.946	-0.018	428.25	0.000
		3	0.919	-0.014	625.89	0.000
		4	0.892	-0.007	813.01	0.000
		5	0.866	-0.002	990.13	0.000
		6	0.841	0.004	1157.9	0.000
		7	0.817	0.007	1316.8	0.000
		8	0.794	0.007	1467.7	0.000
		9	0.772	0.007	1611.0	0.000
		10	0.751	0.005	1747.2	0.000
		11	0.731	0.003	1876.7	0.000
		12	0.711	0.002	1999.9	0.000
		13	0.692	0.002	2117.3	0.000
		14	0.674	0.001	2228.9	0.000
		15	0.656	0.000	2335.2	0.000

图3.4 原序列自相关和偏相关函数图

自相关函数没有快速衰减，可见为非平稳序列。一阶差分后（差分可以使用Matlab函数的diff实现）的自相关和偏相关函数如图3.5所示：

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.952	0.952	210.21	0.000
		2	0.885	-0.229	392.55	0.000
		3	0.797	-0.219	541.35	0.000
		4	0.696	-0.148	655.19	0.000
		5	0.589	-0.051	737.18	0.000
		6	0.486	0.016	793.15	0.000
		7	0.388	-0.003	828.95	0.000
		8	0.302	0.051	850.83	0.000
		9	0.239	0.129	864.57	0.000
		10	0.192	0.030	873.47	0.000
		11	0.154	-0.074	879.20	0.000
		12	0.123	-0.066	882.86	0.000
		13	0.093	-0.080	884.98	0.000
		14	0.065	-0.024	886.02	0.000
		15	0.045	0.078	886.52	0.000

图3.5 原序列一阶差分后自相关和偏相关函数图

自相关函数仍然没有快速衰减，所以一阶差分后的序列依然为非平稳序列，进行二阶差分（一般不超过二阶）后的自相关和偏相关函数如图3.6所示：

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.523	0.523	63.363	0.000
		2 0.197	-0.105	72.392	0.000
		3 0.127	0.095	76.181	0.000
		4 0.004	-0.124	76.185	0.000
		5 -0.056	-0.005	76.925	0.000
		6 -0.051	-0.017	77.549	0.000
		7 -0.170	-0.180	84.426	0.000
		8 -0.260	-0.118	100.63	0.000
		9 -0.171	0.038	107.65	0.000
		10 -0.082	0.019	109.26	0.000
		11 -0.036	0.010	109.62	0.000
		12 0.024	0.025	109.77	0.000
		13 -0.008	-0.076	109.78	0.000
		14 -0.064	-0.061	110.79	0.000
		15 -0.076	-0.085	112.23	0.000

图3.6 原序列二阶差分后自相关和偏相关函数图

自相关系数和偏自相关系数都衰减, 序列趋于平稳。同时自相关系数为正弦波衰减(拖尾), 而偏自相关系数具有截尾性, 所以序列可以使用AR模型建模。

2、确定模型阶次和参数

使用第2章所述模型阶次确定的方法, 确定模型的阶次(Matlab中可以使用arorder、maorder命令估计阶次)。确定模型类型、阶次后, 对于AR模型可以使用自相关法、Burg算法、协方差、改进的协方差法等求出模型参数。本文采用Burg算法。

3、模型拟合

模型类型、阶次、参数确定后, 就可以建立相应的模型对序列进行拟合(如图3.7所示), 检验模型好坏。

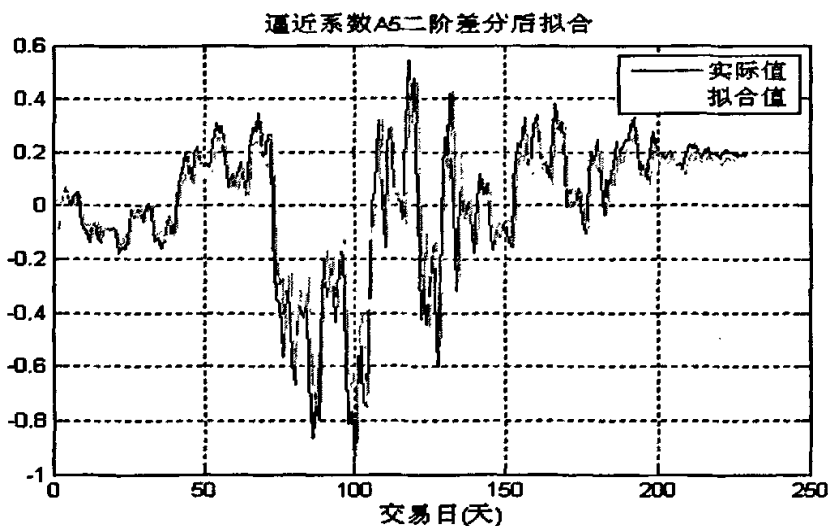


图3.7 逼近系数二阶差分模型曲线拟合图

3.3.4 细节系数建模

对D1~D5细节系数分析后发现细节系统基本都具有平稳性, 因此可以直接建立ARMA模型。通过第2章所述方法, 确定各自曲线的模型类型、阶次和参数。得到的各模型与对应细节系数重构信号的拟合图如图3.8所示。

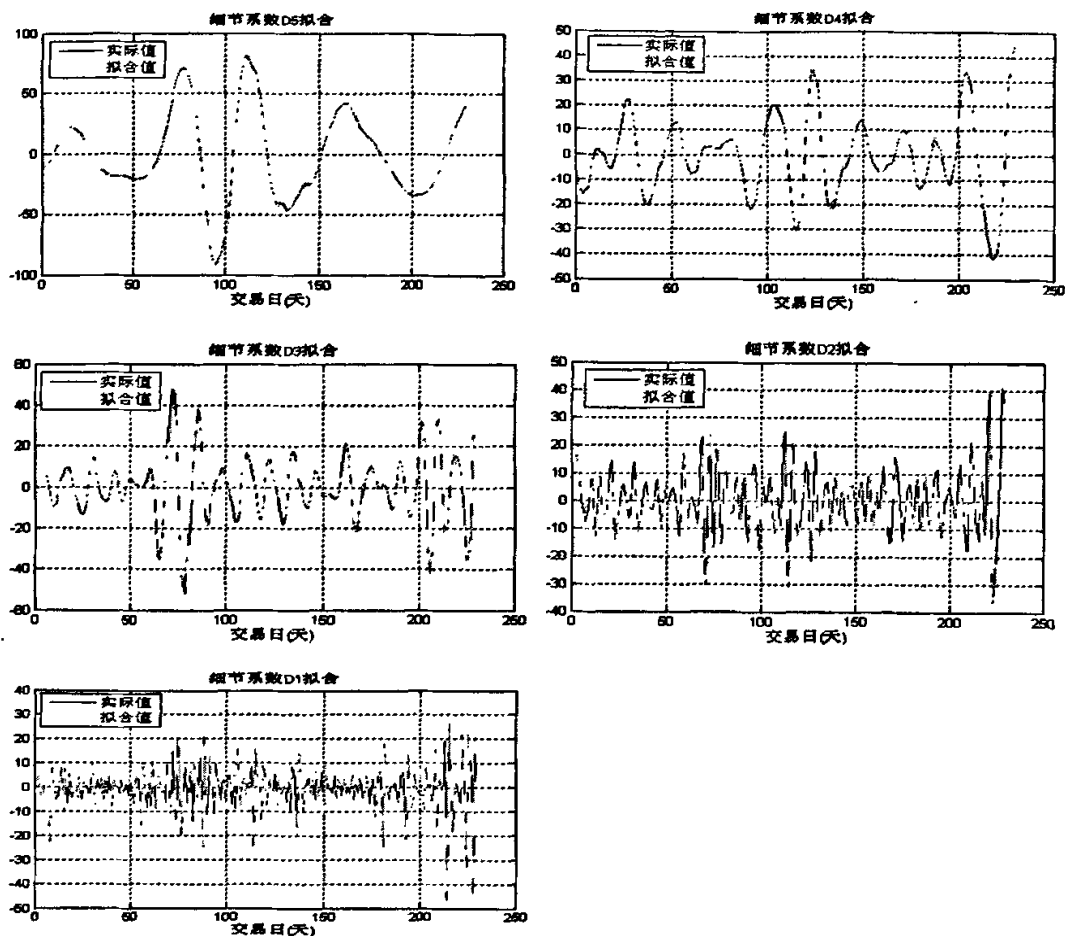


图3.8 各细节系数重构信号模型曲线拟合图

3.3.5 通过模型预测

对各模型按照公式(3.5)进行外推预测, 得到预测值 $\hat{d}_{1,M+k}, \hat{d}_{2,M+k}, \dots, \hat{d}_{J,M+k}, \hat{C}_{J,M+k}$, 则由(3.6)式得到原序列 x_t 在 M 时刻 k 步的预测值为:

$$\hat{x}_{M+k} = \hat{d}_{1,M+k} + \hat{d}_{2,M+k} + \dots + \hat{d}_{J,M+k} + \hat{C}_{J,M+k}$$

预测结果与实际数据的比较如表3.4所示(从2007.1.4到2007.1.31, 共20个交易日的原始数据与预测值的比较)。

表3.4 小波分析预测法预测值和实际值

日期	实际值	预测值	误差	日期	实际值	预测值	误差
2007-1-04	2715.72	2680.36	35.36	2007-1-18	2756.98	2741.32	15.66
2007-1-05	2641.33	2662.79	-21.46	2007-1-19	2832.21	2794.91	37.30
2007-1-08	2707.20	2524.00	183.20	2007-1-22	2933.19	2785.11	148.08
2007-1-09	2807.80	2829.17	-21.37	2007-1-23	2949.14	2830.87	118.27

2007-1-10	2825.58	2764.91	60.67	2007-1-24	2975.13	2857.12	118.01
2007-1-11	2770.11	2590.35	179.76	2007-1-25	2857.36	2833.64	23.72
2007-1-12	2668.11	2549.11	119.00	2007-1-26	2882.56	2820.54	62.02
2007-1-15	2794.70	2712.24	82.46	2007-1-29	2945.26	2880.55	64.71
2007-1-16	2821.02	2773.03	47.99	2007-1-30	2930.56	2940.12	-9.56
2007-1-17	2778.90	2694.36	84.54	2007-1-31	2786.33	2929.53	-143.20

均方根值: $S_N(t) = 95.6065$

多步预测曲线与实际数据曲线的比较如图3.9所示。

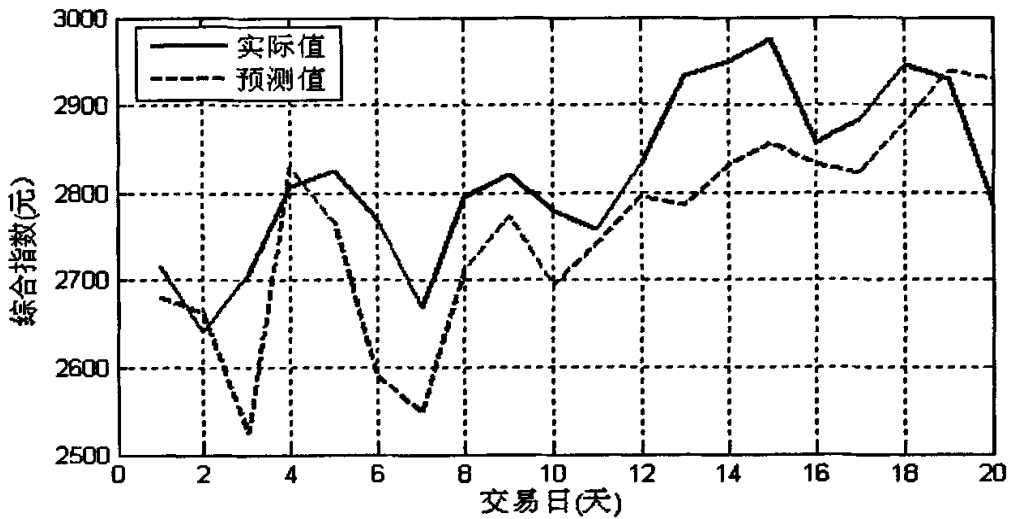


图3.9 小波分析预测法预测曲线和实际曲线比较图

如果不使用小波分析,则原序列趋势项和周期项、随机项混在一起,在差分处理中,由于周期项、随机项的干扰,往往很难得出正确的预测结果。

图3.10所示为实际数据直接采用 $ARIMA(4,2,4)$ 建模后的外推预测图。

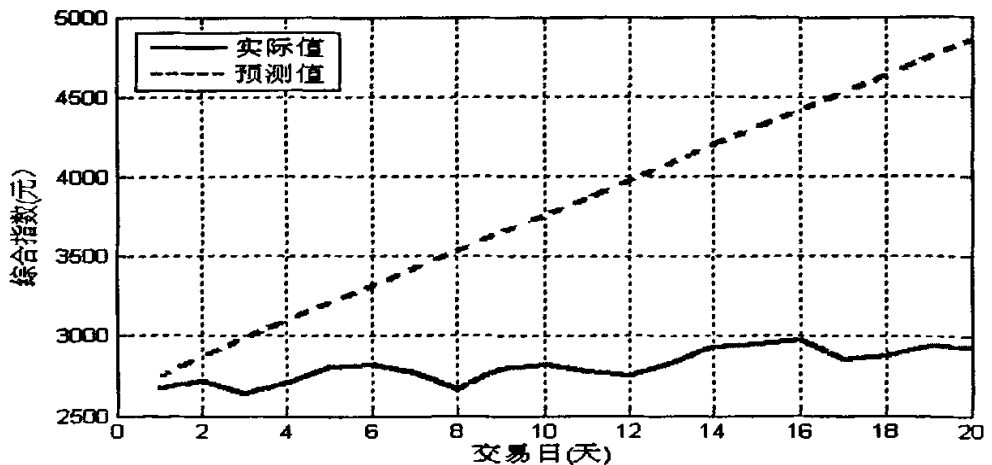


图3.10 $ARIMA(4,2,4)$ 模型预测曲线和实际曲线比较图

§4 基于Kalman滤波的非平稳时间序列建模与预测

Kalman 滤波方法是一种具有严格数学理论基础的预报—校正统计优化方法, 这种优化的特点是能够不断地将观测资料同化到动态系统中, 以数值模式为动力约束条件, 在使观测和模式结果误差方差达到最小的条件下, 得到基于以前所有观测和当前观测的系统变量的最优线性估计 (参见文献[8])。 *Kalman* 滤波方法的另一个特点是它的递推特性, 可以借助于前一时刻的滤波结果递推出当前时刻的状态估计值, 从而大大减少计算量和存储量。

4.1 卡尔曼滤波器

4.1.1 状态方程和量测方程

随机信号 $x(k)$ 的模型方程为:

$$x(k) = A(k)x(k-1) + w(k-1)$$

为了书写方便, 将变量 k 放在下标表示, 则上式成为

$$x_k = Ax_{k-1} + w_{k-1} \quad (4.1)$$

它的量测方程为:

$$y_k = Cx_k + v_k \quad (4.2)$$

(4.1) 中, A 称为状态转移矩阵, $w(k)$ 是信号模型中的白噪声激励。(4.2) 中, C 称为量测矩阵, 它是一个 $m \times n$ 的矩阵 (m 为 y_k 的维数, n 为 x_k 的维数), $v(k)$ 是信号传输或测量中引入的白噪声。

4.1.2 滤波过程方程

卡尔曼滤波要解决的问题是要寻找在最小均方误差下 x_k 的估计值 \hat{x}_k 。具体讲, 设已知动态系统的状态方程和量测方程分别为 (4.1), (4.2) 所示, 且 A 和 C 是已知的, y_k 是测量到的数据也为已知, 求在最小均方误差下 x_k 的估计值 \hat{x}_k 。

假设 w_k 和 v_k 都是均值为零的正态白噪声, 且 w_k 和 v_k 互不相关, 即 $E[w_k] = 0$, $E[v_k] = 0$, $\text{cov}[w_k, v_j] = E[w_k v_j^T] = 0 \quad (k, j = 0, 1, 2, \dots)$

有

$$\text{cov}[w_k, w_j] = E[w_k w_j^T] = Q_k \delta_{kj} \quad (4.3)$$

式中 $Q_k = \text{var}[w_k] = E[w_k w_k^T]$, 是动态噪声的方差阵;

$$\text{cov}[v_k, v_j] = E[v_k v_j^T] = R_k \delta_{kj} \quad (4.4)$$

式中 $R_k = \text{var}[v_k] = E[v_k v_k^T]$, 为统计误差的方差阵。

我们可以得到卡尔曼一步递推公式(具体推导过程略):

$$K_k = P_k' C^T (C P_k' C^T + R_k)^{-1} \quad (4.5)$$

$$P_k' = A P_{k-1} A^T + Q_{k-1} \quad (4.6)$$

$$P_k = P_k' - K_k C P_k' = (I - K_k C) P_k' \quad (4.7)$$

最优估计值为

$$\hat{x}_k = A \hat{x}_{k-1} + K_k (y_k - C A \hat{x}_{k-1}) \quad (4.8)$$

式中, K_k 称为卡尔曼增益, P_k 和 P_k' 为预测误差方差阵和滤波误差方差阵。当我们已知 K_k , 利用前一个 x_k 的最优估计值 \hat{x}_{k-1} 和当前的量测值 y_k 可以求得 \hat{x}_k 。故根据已知的矩阵 A 、 C 、 Q_k 、 R_k 以及观察到的 y_k , 就能用递推算法得到所有的 \hat{x}_1 、 \hat{x}_2 、 \dots 、 \hat{x}_k 以及 P_1 、 P_2 、 \dots 、 P_k 。

4.1.3 外推预测

在纯预测情况下, $w_k = 0$, 同时, v_k 可不考虑, 于是有递推公式:

$$\hat{x}_{k+1} = A \hat{x}_k$$

$$\hat{x}_{k+2} = A(A \hat{x}_k) = A^2 \hat{x}_k$$

...

$$\hat{x}_{k+l} = A^l \hat{x}_k \quad (4.9)$$

则预测值为

$$\hat{y}_{k+l} = C \hat{x}_{k+l} = C A^l \hat{x}_k \quad (4.10)$$

4.2 Kalman 滤波在上证综合指数预测中的应用

4.2.1 模型确定

在进行上证综合指数预测时, 假设指数变量包含两个状态: 当前的指数价格和当前的指数趋势, 分别以 s 和 d 表示, 则

$$x_k = [s_k, d_k]^T, \quad s_{k+1} = s_k + d_k$$

假设每日的变化趋势相同, 即 $d_{k+1} = d_k$, 则 $\begin{bmatrix} s_{k+1} = s_k + d_k \\ s_{k+1} = s_k \end{bmatrix}$, 由于状态的变化很复杂, 并不能完全由指数价格和指数变化趋势准确描述, 则噪声 w_{k-1} 引入到等式中, 表示在综合指数变化趋势线上的“起伏”, 这是符合实际和更一般的情况, 于是有

$$x_k = A x_{k-1} + w_{k-1}$$

也就是(4.1)式, 且状态转移矩阵 $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ 。

因为综合指数可以被直接测量, 而 y_k 是一个只有一行一列的矢量, 并且观察和统计会有一些误差, 因此噪声 v_k 引入到等式中, 于是有

$$y_k = Cx_k + v_k$$

也就是(4.2)式, 且观察矩阵 $C = [1 \ 0]$ 。

至此我们得到了指数预测以及进行递推计算的全部理论公式, 包括:

$$x_k = Ax_{k-1} + w_{k-1} \quad (4.1)$$

$$y_k = Cx_k + v_k \quad (4.2)$$

$$K_k = P_k' C^T (C P_k' C^T + R_k)^{-1} \quad (4.5)$$

$$P_k' = A P_{k-1} A^T + Q_{k-1} \quad (4.6)$$

$$P_k = P_k' - K_k C P_k' = (I - K_k C) P_k' \quad (4.7)$$

$$\hat{x}_k = A \hat{x}_{k-1} + K_k (y_k - C A \hat{x}_{k-1}) \quad (4.8)$$

系统的预测结果为:

$$\hat{y}_{k+1} = C \hat{x}_{k+1} = C A' \hat{x}_k \quad (4.10)$$

4.2.2 模型应用

我们以上证综合指数为例进行 *Kalman* 滤波与预测过程。将样本数据分成估计样本与评价样本两部分。估计样本(从2006.1.4 到2006.12.29 共229个交易日)用于对模型的参数估值, 而评价样本(2007.1.4 到2007.1.31 共20个交易日)用来对模型的预测结果进行检验。

依照前述公式我们得到 x_k 的最优估计值 \hat{x}_k , 再得到一步预测结果 \hat{y}_{k+1} , \hat{y}_{k+1} 与 y_{k+1} 的拟合图如图4.1所示:

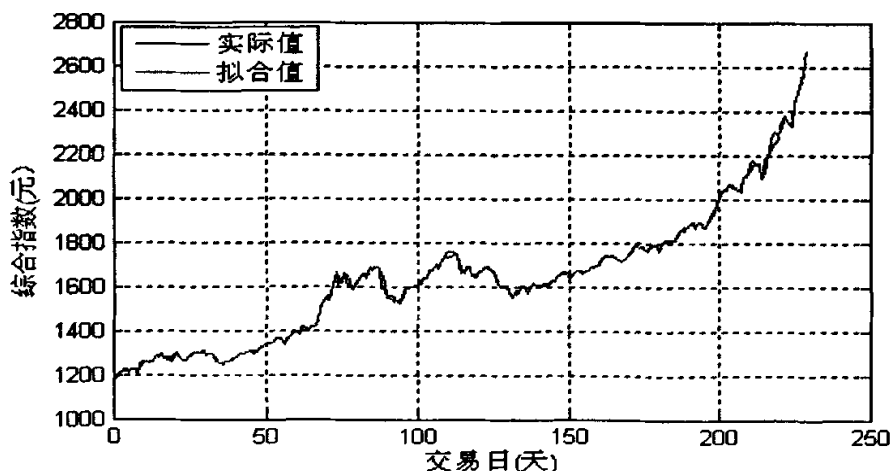


图4.1 Kalman滤波曲线拟合图

进一步采用公式(4.10)可以得到外推20步的预测结果 $\hat{y}_{k+1}, \hat{y}_{k+2}, \dots, \hat{y}_{k+20}$, 与实际数据的比较如表4.1所示(从2007.1.4到2007.1.31, 共20个交易日的原始数据与预测值的比较):

表4.1 Kalman预测值与实际值

日期	实际值	预测值	误差	日期	实际值	预测值	误差
2007-1-04	2715.72	2668.44	47.28	2007-1-18	2756.98	3172.21	-415.23
2007-1-05	2641.33	2718.81	-77.48	2007-1-19	2832.21	3222.58	-390.37
2007-1-08	2707.20	2769.19	-61.99	2007-1-22	2933.19	3272.96	-339.77
2007-1-09	2807.80	2819.57	-11.77	2007-1-23	2949.14	3323.34	-374.20
2007-1-10	2825.58	2869.94	-44.36	2007-1-24	2975.13	3373.71	-398.58
2007-1-11	2770.11	2920.32	-150.21	2007-1-25	2857.36	3424.09	-566.73
2007-1-12	2668.11	2970.70	-302.59	2007-1-26	2882.56	3474.47	-591.91
2007-1-15	2794.70	3021.08	-226.38	2007-1-29	2945.26	3524.84	-579.58
2007-1-16	2821.02	3071.45	-250.43	2007-1-30	2930.56	3575.22	-644.66
2007-1-17	2778.90	3121.83	-342.93	2007-1-31	2786.33	3625.60	-839.27

均方根值： $S_N(t) = 401.6479$

多步预测曲线与实际数据曲线的比较如图4.2所示。

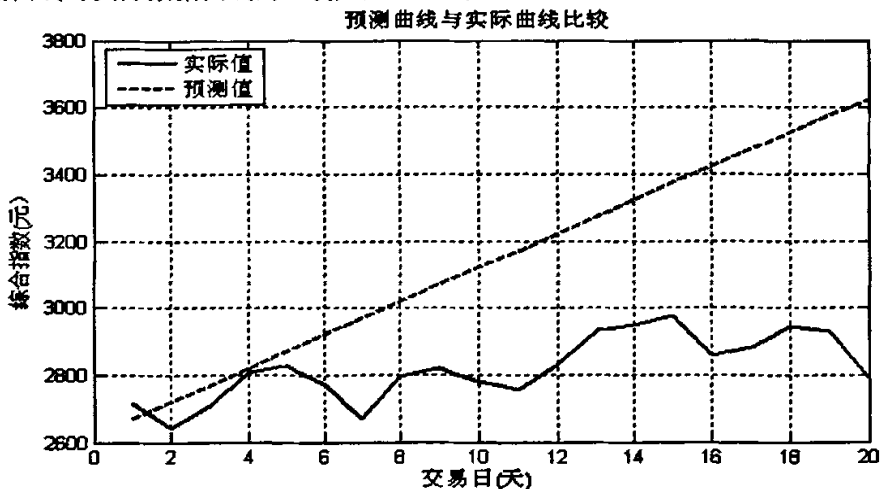


图4.2 Kalman外推预测曲线与实际曲线比较图

通过上述过程发现：在一步预测时，由于不断有新息加入，曲线得到了很好的修正，而多步外推预测时，由于没有新息加入，各参数维持在最近的修正值上，使预测误差越来越大，同时 Kalman 外推预测是不收敛的，说明 Kalman 滤波只适用于短期预测，随着预测长度的增加，预测误差也会增大。

4.3 Kalman-ARMA 预测方法

4.3.1 概述

针对上述 Kalman 外推预测误差增大的情况，将 $ARMA(p,q)$ 模型法引入，利用

$ARMA(p, q)$ 模型预测的收敛性, 可以提高 *Kalman* 预测精度。

如前述, 指数变量包含两个状态: 当前的指数价格和当前的指数趋势, 分别以 s 和 d 表示, 即 $x_k = [s_k, d_k]^T$, 通过 *Kalman* 滤波过程得到了 x_k 的最优估计值 \hat{x}_k , 相应的也得到了 \hat{d}_k 。分析 \hat{d}_k 后发现 \hat{d}_k 可近似看作平稳序列。对 $\hat{d}_1, \hat{d}_2, \dots, \hat{d}_k$ 建立 $ARMA(p, q)$ 模型, 并依据模型可以得到预测值 $\hat{d}_{k+1}, \hat{d}_{k+2}, \dots, \hat{d}_{k+\tau}$ 。

在 k 时刻外推 l 步预测 \hat{y}_{k+l} 时, 有

$$\begin{aligned}\hat{x}_{k+1} &= \hat{x}_k + \hat{d}_k \\ \hat{x}_{k+2} &= \hat{x}_{k+1} + \hat{d}_{k+1} = \hat{x}_k + \hat{d}_k + \hat{d}_{k+1} \\ \hat{x}_{k+3} &= \hat{x}_{k+2} + \hat{d}_{k+2} = \hat{x}_k + \hat{d}_k + \hat{d}_{k+1} + \hat{d}_{k+2} \\ &\dots \\ \hat{x}_{k+\tau} &= \hat{x}_k + \sum_{i=1}^l \hat{d}_{k+i}\end{aligned}\quad (4.11)$$

则预测值

$$\hat{y}_{k+l} = C\hat{x}_{k+l}\quad (4.12)$$

4.3.2 Kalman-ARMA 法在上证综合指数预测中的应用

将此方法应用到上证综合指数的预测中。首先对上证综合指数进行 *Kalman* 滤波过程, 得到最优估计 $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_k$, 从 $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_k$ 中提取 $\hat{d}_1, \hat{d}_2, \dots, \hat{d}_k$ 得到趋势曲线如图 4.3 所示:

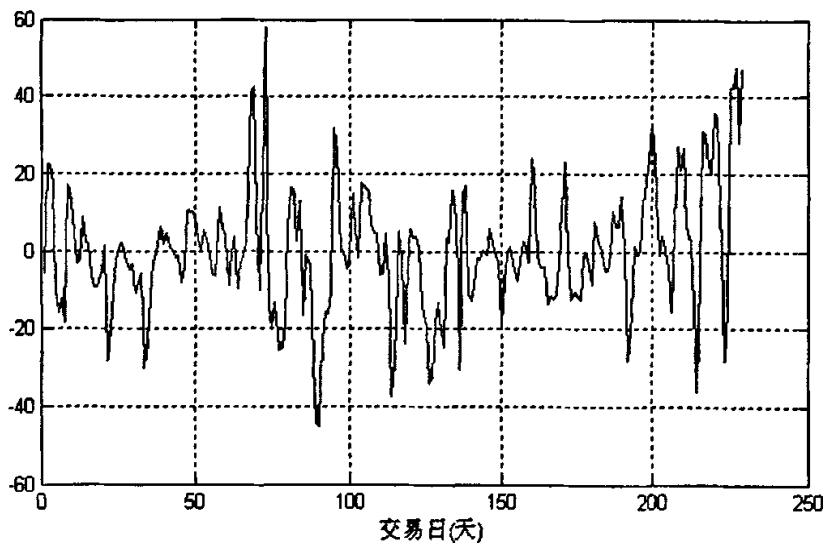


图 4.3 上证综合指数趋势曲线

可发现此序列可近似看作均值为 0 的平稳过程, 建立此过程的 $ARMA(p, q)$ 模型。模型与 $\hat{d}_1, \hat{d}_2, \dots, \hat{d}_k$ 趋势曲线的拟合图如图 4.4 所示。

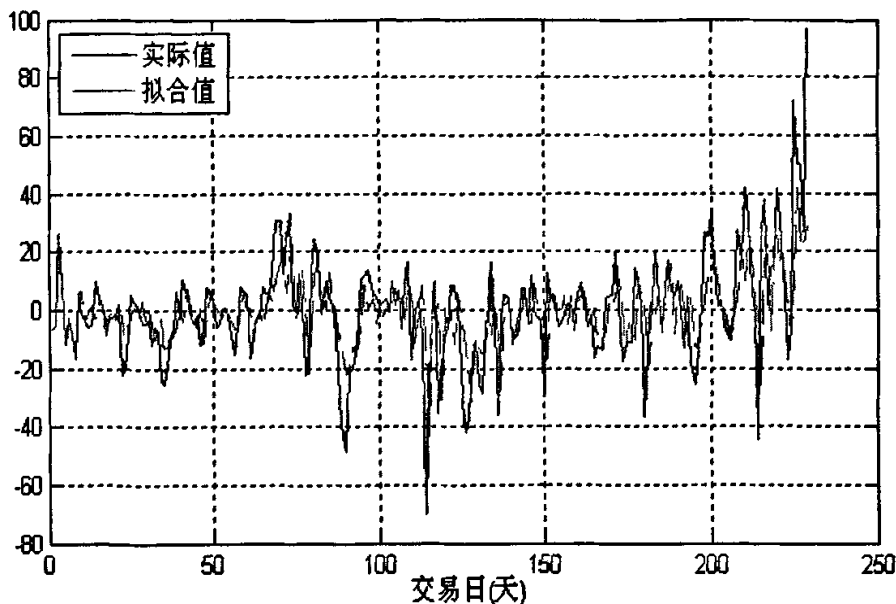


图4.4 上证综合指数趋势曲线拟合图

进一步利用 $ARMA(p,q)$ 模型外推预测 $\hat{d}_{k+1}, \hat{d}_{k+2}, \dots, \hat{d}_{k+20}$ ，根据公式 (4.11), (4.12) 可以得到预测值 $\hat{y}_{k+1}, \hat{y}_{k+2}, \dots, \hat{y}_{k+20}$ ，预测结果与实际数据的比较如表4.2所示（从2007.1.4到2007.1.31，共20个交易日的原始数据与预测值的比较）：

表4.2 Kalman-ARMA(p,q)预测数据与实际数据

日期	实际值	预测值	误差	日期	实际值	预测值	误差
2007-1-04	2715.72	2672.96	42.76	2007-1-18	2756.98	2855.35	-98.37
2007-1-05	2641.33	2710.96	-69.63	2007-1-19	2832.21	2864.14	-31.93
2007-1-08	2707.20	2741.22	-34.02	2007-1-22	2933.19	2873.69	59.50
2007-1-09	2807.80	2765.96	41.84	2007-1-23	2949.14	2882.17	66.97
2007-1-10	2825.58	2788.41	37.17	2007-1-24	2975.13	2889.57	85.56
2007-1-11	2770.11	2809.55	-39.44	2007-1-25	2857.36	2894.20	-36.84
2007-1-12	2668.11	2817.02	-148.91	2007-1-26	2882.56	2896.74	-14.18
2007-1-15	2794.70	2824.62	-29.92	2007-1-29	2945.26	2899.31	45.95
2007-1-16	2821.02	2834.53	-13.51	2007-1-30	2930.56	2901.76	28.80
2007-1-17	2778.90	2844.98	-66.08	2007-1-31	2786.33	2904.17	-117.84

均方根值： $S_N(t) = 65.0260$

多步预测曲线与实际数据曲线的比较如图4.5所示。

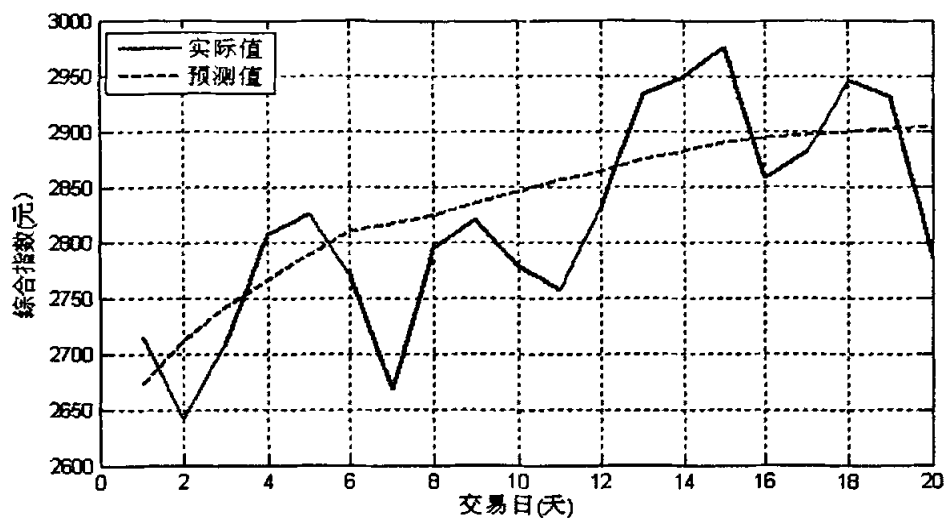


图4.5 Kalman-ARMA预测曲线与实际曲线比较图

通过上述过程可以发现：*Kalman-ARMA*方法提高了预测精度，克服了*Kalman*预测不收敛的缺陷，降低了预测误差。

§5 基于灰度组合模型的非平稳时间序列建模与预测

5.1 概述

由邓聚龙教授建立的灰色预测模型 $GM(1,1)$ ，由于具有所需原始数据不多，容易采集，方法简便并具有较高的准确性，在近二十年来得到了广泛的应用（参见文献[23]）。它只需很少几个数据就可以建立模型进行预测，尤其对短期预测具有较高精度，但因为该预测模型是指数函数形式的，具有单调性，所以进行长期预测时，预测值往往偏高或偏低，因而对随机波动性较大的数据拟合较差，预测精度低^[24]。

而 $ARMA(p,q)$ 模型对于长期预测具有较好拟合效果，但它分析的对象为平稳时间序列，对于一般的时间序列是很难满足的。将灰色理论和 $ARMA(p,q)$ 模型联合应用，利用 $GM(1,1)$ 模型提取非平稳时间序列中的趋势项，使非平稳时间序列转化为平稳时间序列，来满足 $ARMA(p,q)$ 模型对时序平稳、零均值和正态的要求。利用常规的 $ARMA(p,q)$ 模型对提取趋势项后的残差部分进行时序分析，建模并预测，获取时序预测修正数据，将此时序修正数据与 $GM(1,1)$ 进行叠加，即可获得更高预测精度。从 $GM(1,1)$ 模型预测的角度分析，该方法与单纯的灰色理论预测不同，在 $GM(1,1)$ 模型预测的基础上通过残差时序分析又进行了一次修正，进一步提高精度。从时间序列预测的角度分析，该方法通过 $GM(1,1)$ 模型来提取非平稳时序的趋势项比传统方法相比具有更高的精度，能克服传统方法的固有缺点^[24]。

5.2 灰色 $GM(1,1)$ -ARMA 组合模型

5.2.1 组合模型

对于非平稳时间序列 $\{x_t\}$ ($t=1,2,\dots,N$)，是确定性部分与平稳随机部分的叠加。因此，可采用确定性函数来描述其确定性部分，采用 $ARMA(p,q)$ 模型来描述其平稳随机部分，再将确定性函数与 $ARMA(p,q)$ 模型组合成为平稳时序模型。记确定性部分为 d_t ，平稳随机部分为 s_t ，则

$$x_t = d_t + s_t \quad (5.1)$$

式中 d_t 称为趋势项，是随时间 t 变化的某一确定性函数，如线形函数、幂函数、指数函数、周期函数等，这里采用 $GM(1,1)$ 模型提取趋势项 d_t ； s_t 则是从 x_t 中提取了 d_t 后剩下的平稳

随机项, 序列 $\{s_t\}$ 可用 $ARMA(p, q)$ 模型来描述, 将这两部分组合即得到组合模型^[24]。

5.2.2 灰色 GM(1,1)模型提取趋势项

从灰色系统理论的观点看, 随机变量可视为在一定范围内变化的灰色量, 而随机过程可视为在一定范围内变化的灰色过程。灰色系统理论认为: 对一个预测系统, 必然是具有已知信息, 又有未知信息或不确定信息, 且处于被动变化之中。尽管系统的行为现象是朦胧的, 数据是杂乱的, 但它毕竟是有序的, 在这杂乱无章的数据背后, 必然潜在着某种规律。

由于时间序列变量自身变化的复杂性, 受诸多因素制约, 具有很大的不确定性, 所以时间序列变量系统就是一个处于动态变化之中的灰色系统, 可用 $GM(1,1)$ 建模^[24]。

$GM(1,1)$ 模型是单序列的一阶线性动态模型, 设某原始数据系列为:

$$x^{(0)} = [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)] \quad (5.2)$$

做一次累加生成1-AGO为:

$$x^{(1)} = [x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)] \quad (5.3)$$

式中

$$x^{(1)}(t) = \sum_{i=1}^t x^{(0)}(i) = x^{(0)}(t) + x^{(1)}(t-1) \quad t = 1, 2, 3, \dots, n \quad (5.4)$$

对 $x^{(1)}$ 可建立下述白化形式的方程

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \quad (5.5)$$

这是一个变量的一阶微分方程, 故记为 $GM(1,1)$, 其中 a, u 为模型参数。记参数列为 \hat{a} , 则 $\hat{a} = \begin{bmatrix} a \\ u \end{bmatrix}$, 按最小二乘法解, 求 \hat{a} ,

$$\hat{a} = \begin{bmatrix} a \\ u \end{bmatrix} = (B^T B)^{-1} B^T y_n \quad (5.6)$$

$$B = \begin{bmatrix} -\frac{1}{2}(x_1^{(1)}(1) + x_1^{(1)}(2)) & 1 \\ -\frac{1}{2}(x_1^{(1)}(2) + x_1^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x_1^{(1)}(n-1) + x_1^{(1)}(n)) & 1 \end{bmatrix} \quad (5.7)$$

$$y_n = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T \quad (5.8)$$

根据微分方程理论, 白化形式微分方程的解为:

$$x^{(1)}(t+1) = \left(x^{(0)}(1) - \frac{u}{a} \right) e^{-at} + \frac{u}{a} \quad (5.9)$$

对上式做累减还原处理(LAGO), 可得到原始序列 $\{x_t\}$ 的预测公式为:

$$x^{(0)}(t+1) = (-ax^{(0)}(1) + u)e^{-at} \quad (5.10)$$

上式描述了 $\{x_t\}$ 随时间 t 按变动的趋势 d_t 。因此, 根据前面的思路, 也就可以用

$GM(1,1)$ 来提取非平稳时间序列 $\{x_t\}$ 中的趋势项 d_t ,

$$d_t = (-ax^{(0)}(1) + u)e^{-a(t-1)} \quad (5.11)$$

5.2.3 随机项 ARMA 建模

时间序列在分离趋势项后, 剩余的残差部分 s_t 具有平稳随机过程的性质, 因此称为随机项, 可以用平稳随机过程的方法来分析。

根据第2章所述 $ARMA(p, q)$ 模型建立的方法, 通过选择模型类型、确定模型阶次、估计模型参数等步骤建立随机项的模型。

$$s_t = -\sum_{k=1}^p a_k s_{t-k} + \sum_{k=0}^q b_k u(n-k) \quad (5.12)$$

5.2.4 组合模型预测

由 (5.11), (5.12) 得 \hat{x}_t 的 $GM(1,1)-ARMA$ 组合模型为:

$$\hat{x}_t = d_t + s_t = (-ax^{(0)}(1) + u)e^{-a(t-1)} - \sum_{k=1}^p a_k s_{t-k} + \sum_{k=0}^q b_k u(n-k) \quad (5.13)$$

此式等号右边第一项称为 $GM(1,1)$ 部分, 由 $GM(1,1)$ 建模所确定; 其余二项为 $ARMA(p, q)$ 部分, 由随机项的 $ARMA(p, q)$ 模型得到。

在 t 时刻做 \hat{x}_t 的 l 步的外推预测时, 先由公式 (5.11) 和 (2.11) 得到 d_t , s_t 的预测值 \hat{d}_{t+l} , \hat{s}_{t+l} ,

$$\hat{d}_{t+l} = (-ax^{(0)}(1) + u)e^{-a(t+l-1)} \quad (5.14)$$

$$\hat{s}_{t+l} = \begin{cases} -\sum_{k=1}^p a_k s_{t+l-k} & l=1 \\ -\sum_{k=1}^{l-1} a_k \hat{s}_{t+l-k} - \sum_{k=l}^p a_k s_{t+l-k} & 1 \leq l \leq p \\ -\sum_{k=1}^p a_k \hat{s}_{t+l-k} & l > p \end{cases} \quad (5.15)$$

则 \hat{x}_t 的 l 步预测值 \hat{x}_{t+l} 为:

$$\hat{x}_{t+1} = \hat{d}_{t+1} + \hat{s}_{t+1} \quad (4.16)$$

5.3 GM(1,1)-ARMA 组合模型在上证综合指数预测中的应用

将GM(1,1)-ARMA组合模型应用到上证综合指数的预测中。取上证综合指数从2006.1.4到2006.12.29(共有229个交易日)每交易日的收盘价格作为估计样本,用于对模型的参数估值,取2007.1.4到2007.1.31(共20个交易日)的收盘价格作为评价样本,对模型的预测结果进行检验。

5.3.1 GM(1,1)模型提取趋势项

设2006.1.4-2006.12.29上证综合指数序列为 $\{x_t\}$,利用公式(5.11)提取 $\{x_t\}$ 中的趋势项 $\{d_t\}$,如图5.1所示。

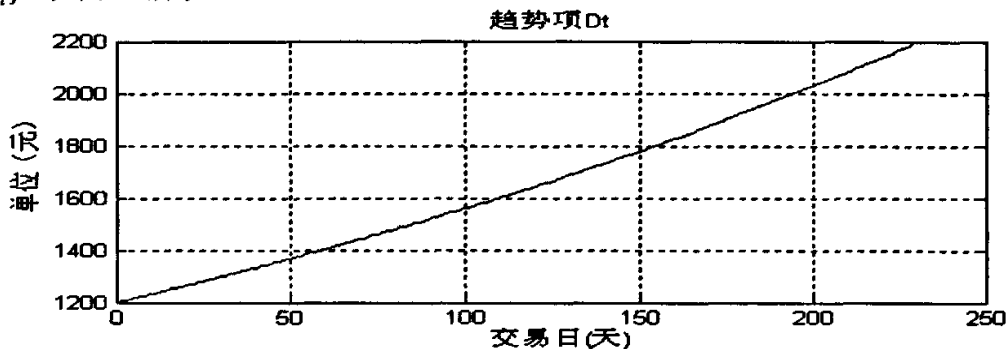


图5.1 GM(1,1)模型提取的趋势项图

5.3.2 随机项建模

由(5.1)得

$$s_t = x_t - d_t \quad (4.17)$$

$\{s_t\}$ 如图5.2所示:

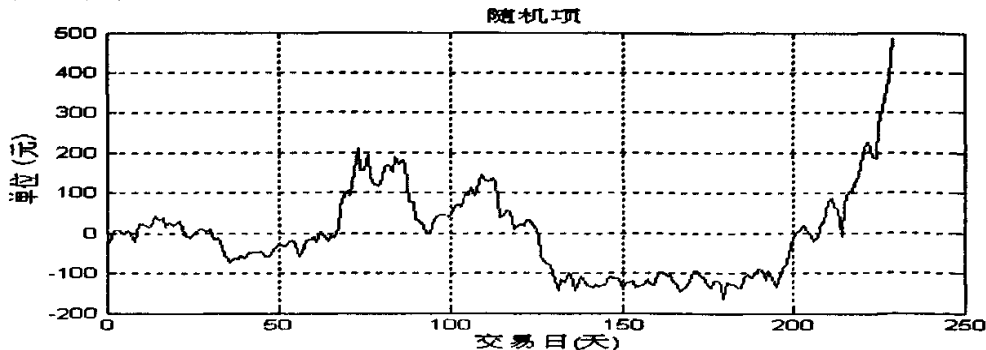


图5.2 随机项 s_t 走势图

对 $\{s_t\}$ 建立 $ARMA(p,q)$ 模型，模型与实际数据的拟合曲线如图5.3所示。

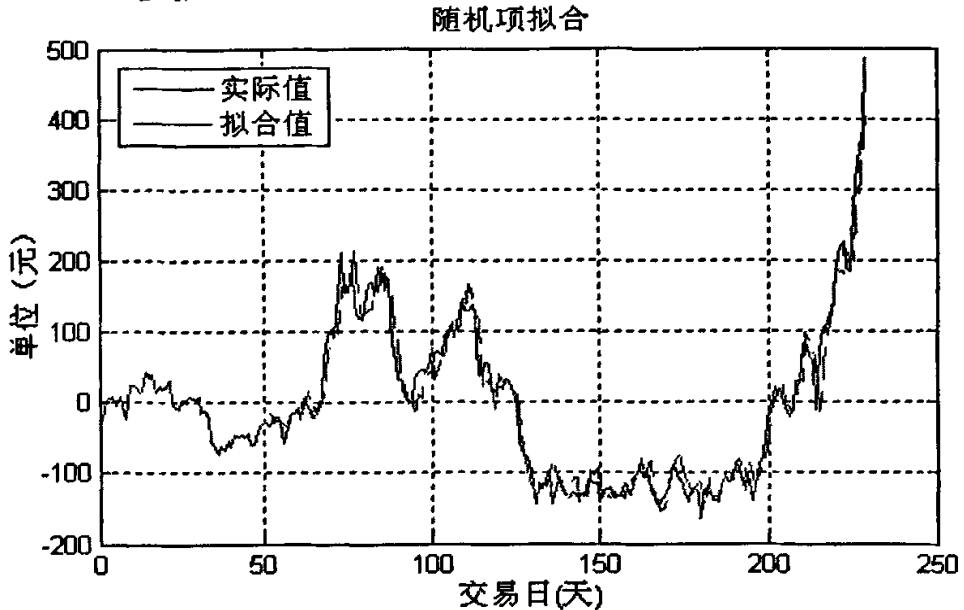


图5.3 随机项模型曲线拟合图

5.3.3 组合模型预测上证指数

使用公式 (5.14)，(5.15) 分别对趋势项和随机项的进行预测，得到预测值 \hat{d}_{t+i} 和 \hat{s}_{t+i} ，则序列 $\{x_t\}$ 的预测值

$$\hat{x}_{t+i} = \hat{d}_{t+i} + \hat{s}_{t+i} \quad (5.14)$$

预测结果与实际数据的比较如表5.1所示（从2007.1.4到2007.1.31，共20个交易日）。

表5.1 $GM(1,1)-ARMA(p,q)$ 组合模型预测法预测值和实际值

日期	实际值	预测值	误差	日期	实际值	预测值	误差
2007-1-04	2715.72	2696.22	19.50	2007-1-18	2756.98	2727.16	29.82
2007-1-05	2641.33	2717.85	-76.52	2007-1-19	2832.21	2700.13	132.08
2007-1-08	2707.20	2729.03	-21.83	2007-1-22	2933.19	2701.88	231.31
2007-1-09	2807.80	2763.57	44.23	2007-1-23	2949.14	2688.34	260.80
2007-1-10	2825.58	2777.97	47.61	2007-1-24	2975.13	2677.67	297.46
2007-1-11	2770.11	2804.97	-34.86	2007-1-25	2857.36	2653.57	203.79
2007-1-12	2668.11	2744.47	-76.36	2007-1-26	2882.56	2638.24	244.32
2007-1-15	2794.70	2737.00	57.70	2007-1-29	2945.26	2628.94	316.32
2007-1-16	2821.02	2728.56	92.46	2007-1-30	2930.56	2614.33	316.23
2007-1-17	2778.90	2740.20	38.70	2007-1-31	2786.33	2601.82	184.51

均方根值: $S_N(t) = 172.4857$

多步预测曲线与实际数据曲线的比较如图5.4所示。

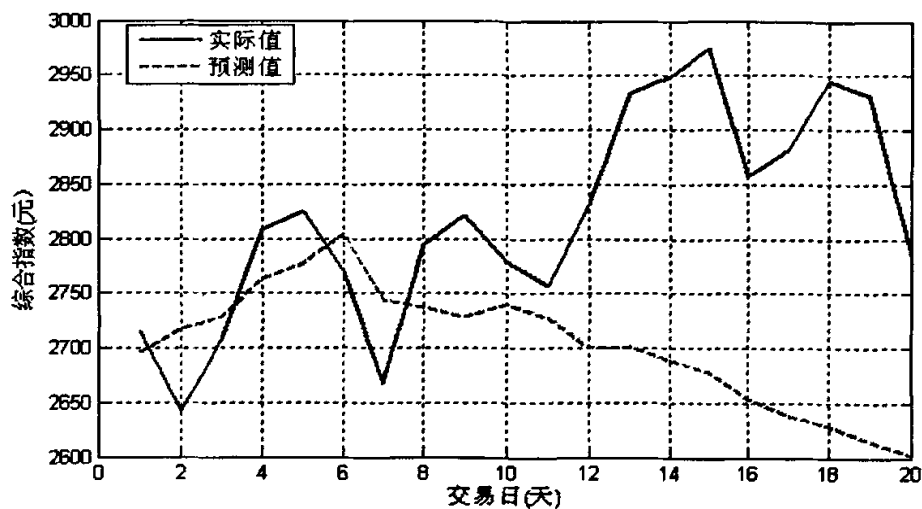


图5.4 GM(1,1)-ARMA组合模型预测曲线和实际曲线比较图

可见随着预测步数的增大预测误差也越来越大，灰度预测模型适于趋势项明显而减去趋势项后的随机项趋于平稳的时间序列分析，但不适于趋势项中蕴含周期项，随机项不平稳的时间序列分析，其结果会使预测误差增大。

§6 BP神经网络预测

神经网络自上世纪90年代再次复兴以来,因为人工神经网络广泛的适应能力、学习能力和映射能力,在多变量非线性系统的建模方面取得了惊人的成就,成为新兴的预测时间序列的方法(参见文献[18])。人工神经网络模型具有巨量并行性,存储分布性,结构可变性,高度非线性,自学习性和自组织等特点,而且可以逼近任何连续函数,目前广泛应用神经网络作为非线性函数逼近模型以及混沌时间序列的预测上。神经网络的最大优点是不依赖于模型,所以比较适合用于非平稳时间序列的预测^[10]。

6.1 概述

6.1.1 神经网络的简单原理

人工神经网络是根据人的认识过程而开发出的一种算法。假如我们现在只有一些输入和相应的输出,而对如何由输入得到输出的机理并不清楚,那么我们可以把输入与输出之间的未知过程看成是一个“网络”,通过不断地给这个网络输入和相应的输出来“训练”这个网络,网络根据输入和输出不断地调节自己的各节点之间的权值来满足输入和输出。这样,当训练结束后,我们给定一个输入,网络便会根据自己已调节好的权值计算出一个输出。这就是神经网络的简单原理。

6.1.2 神经元和神经网络的结构

如上所述,神经网络的基本结构如图 6.1 所示:

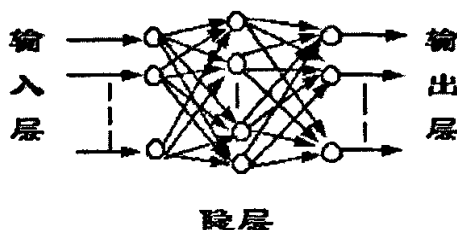


图 6.1 神经网络基本结构图

神经网络一般都有多层,分为输入层,输出层和隐含层,层数越多,计算结果越精确,但所需的时间也就越长,所以实际应用中要根据要求设计网络层数。

神经网络中每一个节点叫做一个人工神经元，他对应于人脑中的神经元，两者的结构比较如图6.2所示：

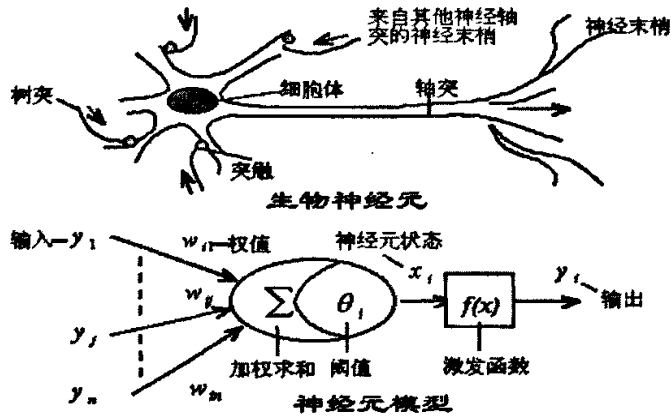


图 6.2 生物神经元与神经元模型比较图

一个人工神经元一般有多个输入和一个输出，另外有一个激发函数，不同的激发函数对应了不同的网络，也决定了网络的用途。

6.1.3 BP 神经网络

1、概述

BP神经网络是应用得最为广泛，最为重要的一种神经网络。BP算法的基本思想是利用最小均方差(LMS)算法，在网络的学习过程中使用梯度搜索技术，利用误差向后传播来修正权值，从而实现网络的实际输出与期望输出的均方差的最小化^[26]。

通常的神经网络结构中包含一个输入层，若干隐含层和一个输出层。图6.3所示是含有一个隐含层的BP神经网络模型。其中输入层神经元的个数为 n_0 ，隐层神经元的个数为 n_1 ，输出层神经元的个数为 n_2 。 $W^{(s)}$ 表示第 s 层的权矩阵， $s=1,2$ 。 $X_{out}^{(s)}$ 表示第 s 层的输出向量。 $V^{(s)}$ 表示第 s 层的净输入向量。

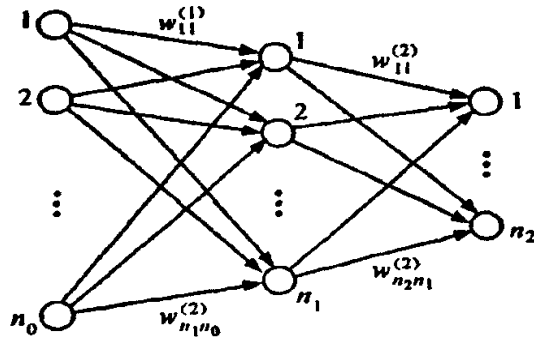


图 6.3 含有一个隐含层的 BP 网络模型图

标准的BP 算法基于最速下降法, 设能量函数如下:

$$E_q = \frac{1}{2} (d_q - X_{out}^{(2)})^T (d_q - X_{out}^{(2)}) = \frac{1}{2} \sum_{h=1}^{n_2} (d_{qh} - x_{out,h}^{(2)})^2 \quad (6.1)$$

其中, d_q 是神经网络对应于(样本)第 q 个输入的期望输出, $X_{out}^{(2)}$ 是网络相应的实际输出。网络的学习过程就是对每一个 $w_{ji}^{(s)}$ 进行调整, 使得能量函数 E_q 随着迭代的进行而减小, 从而最终达到某一个全局的极小值。根据最速下降法可得出关于各层权矩阵的更新迭代式。

$$w_{ji}^{(s)}(k+1) = w_{ji}^{(s)}(k) + \mu^{(s)} \delta_j^{(s)} x_{out,i}^{(s)} \quad (6.2)$$

式中, $\mu^{(s)}$ 为神经网络第 s 层的学习参数; $\delta_j^{(s)}$ 称为局部误差(或局部梯度)。

2、BP网络学习算法的计算步骤^[10]

- (1) 初始化: 置所有的加权系数为最小的随机数;
- (2) 提供训练集: 给出顺序赋值的输入向量 x_1, x_2, \dots, x_N 和期望的输出向量 t_1, t_2, \dots, t_N ;
- (3) 计算实际输出: 计算隐含层、输出层各神经元的输出;
- (4) 计算期望值与实际输出的误差;
- (5) 调整输出层的加权系数 w_{kj} ;
- (6) 调整隐含层的加权系数 w_{ji} ;
- (7) 返回计算(3)步, 直到误差满足要求为止。

3、BP网络的用途

BP网络的用途十分广泛, 主要用于以下方面:

- (1) 函数逼近: 用输入矢量和相应的输出矢量训练一个网络逼近一个函数;
- (2) 模式识别: 用一个特定的输出矢量将它与输入矢量联系起来;
- (3) 分类: 把输入矢量以所定义的合适方式进行分类;
- (4) 数据压缩: 减少输出矢量维数以便于传输或存储。

本文主要使用BP神经网络完成第一个用途。

6.2 BP 神经网络在上证综合指数预测中的应用

应用 *Matlab* 提供的神经网络工具箱可以很方便的建立时间序列的神经网络模型。国内外都有过一些利用神经网络预测市场证券指数的先例。常见的方法和流程是: (1) 确定某种算法的神经网络(比如BP算法, RBF 算法, 小波算法等); (2) 将预测样本前某一段时期内的数据作为训练样本, 训练神经网络; (3) 将预测样本代入训练好的神经网络进行预测, 得出预测结果(参见文献[18], [19], [20])。

下面将BP神经网络模型应用到上证综合指数的预测中。选取2006.1.4 到2006.12.29 (共229个交易日)的收盘价格作为训练样本和预测样本, 2007.1.4 到2007.1.31(共20个交易日)的收盘价格作为评价样本, 用来对模型的预测结果进行检验。

6.2.1 数据归一化处理

在应用神经网络之前, 需要进行数据预处理, 即将从股市得到的原始数据转化为能被神经网络识别的输入数据, 即数据的归一化处理:

$$x'_i = \frac{x_i - 1/2(x_{\max} + x_{\min})}{x_{\max} - x_{\min}} \quad (6.3)$$

其中 x_i 为原始数据, x_{\max} 、 x_{\min} 分别是原始序列中的最大值和最小值。 x'_i 是处理后的数据, 其范围是 $[-0.5, 0.5]$ 。

得到神经网络的输出数据后进行反归一化处理, 得到实际预测值。

$$x_i = x'_i(x_{\max} - x_{\min}) + \frac{1}{2}(x_{\max} + x_{\min}) \quad (6.4)$$

6.2.2 确定网络拓扑结构

在应用神经网络的过程中, 重要的一步就是确定网络的拓扑结构, 即网络的层次及各个层中神经元的个数。理论上已经证明, 一个三层的神经网络可以以任意精度逼近一个非线性连续函数, 因而选择采用一个隐含层。在输入节点个数方面, 选取前5日、10日、15日、20日、25日的移动平均和前一天的股票价格作为输入, 即MAV 5, MAV 10, MAV 15, MAV 20, MAV 25 和 X_{t-1} , 网络的输出为 X_t , 即6输入1输出。在确定隐含层的神经元数目时, 分别选取2~10 之内的自然数进行试验, 根据试验结果决定隐层神经元的数目。通过对上证指数样本数据进行试验, 得到最优隐层神经元个数为3, 即网络的结构是6-3-1。

6.2.3 训练网络

在建立BP神经网络后, 需要选取某一段时期内的数据作为训练样本。这里选用20组数据作为训练样本, 从数据的第170天开始, 选取每天的前5日、10日、15日、20日、25日的移动平均和当天的股票价格作为输入, 20天后的股票价格作为输出, 即以第170~189天的统计数据作为输入, 第190~209天的股票价格作为输出。以此数据训练网络。仿真输出如图6.4所示。

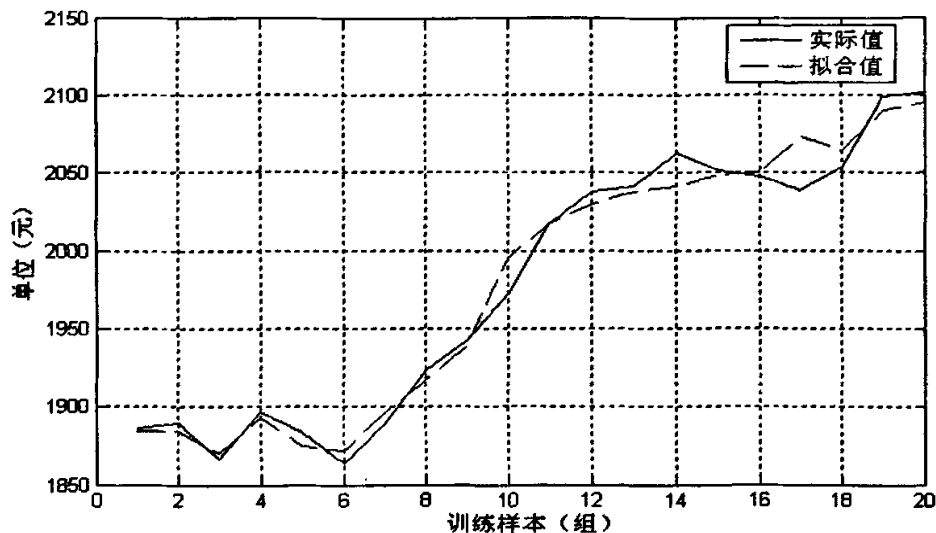


图6.4 BP神经网络仿真输出与实际曲线比较图

6.2.4 预测

将训练好的网络应用到仿真预测中。以第210-229天的统计数据作为输入，预测第230-249天的股票价格(即2007.1.4到2007.1.31)。由于BP神经网络的预测结果并不唯一，因此对训练好的网络进行多次预测，将预测结果取平均后与实际数据的比较如表6.1所示(从2007.1.4到2007.1.31，共20个交易日)。

表6.1 采用BP神经网络得到的预测值和实际值的比较

日期	实际值	预测值	误差	日期	实际值	预测值	误差
2007-1-04	2715.72	2575.61	140.11	2007-1-18	2756.98	2785.51	-28.53
2007-1-05	2641.33	2783.23	-141.90	2007-1-19	2832.21	2787.91	44.30
2007-1-08	2707.20	2771.56	-64.36	2007-1-22	2933.19	2788.87	144.32
2007-1-09	2807.80	2787.28	20.52	2007-1-23	2949.14	2780.09	169.05
2007-1-10	2825.58	2790.68	34.90	2007-1-24	2975.13	2788.21	186.92
2007-1-11	2770.11	2793.92	-23.81	2007-1-25	2857.36	2784.25	73.11
2007-1-12	2668.11	2793.14	-125.03	2007-1-26	2882.56	2789.28	93.28
2007-1-15	2794.70	2769.33	25.37	2007-1-29	2945.26	2783.96	161.30
2007-1-16	2821.02	2726.83	94.19	2007-1-30	2930.56	2724.14	206.42
2007-1-17	2778.90	2743.22	35.68	2007-1-31	2786.33	2660.83	125.50

均方根值： $S_n(t)=113.5959$

预测曲线与实际数据曲线的比较如图6.5所示。

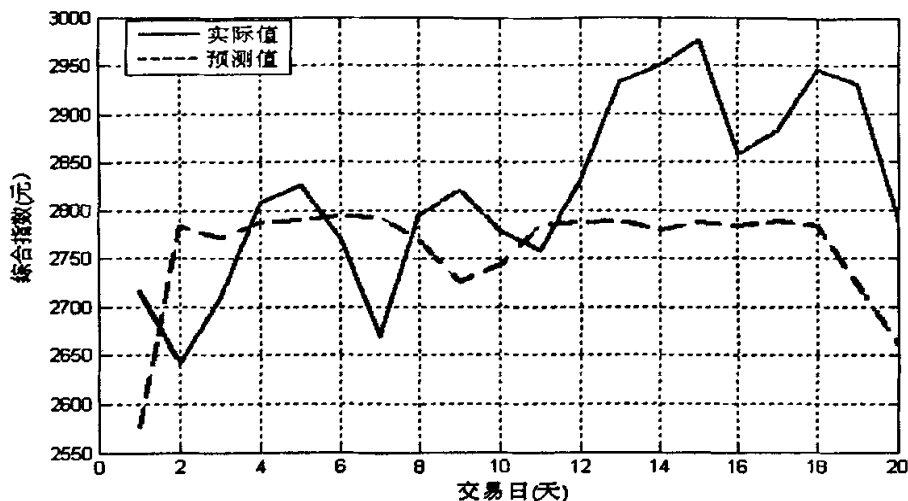


图 6.5 BP 神经网络预测曲线和实际曲线比较图

6.3 BP 神经网络分析

BP神经网络是目前应用最多的一种神经网络形式，但它也不是非常完美的，为了更好的理解应用神经网络进行问题求解，这里对它的优缺点展开讨论^[10]。

6.3.1 BP 神经网络的优点

1、网络实质上实现了一个从输入到输出的映射功能，而数学理论已证明它具有实现任何复杂非线性映射的功能。这使得它特别适合于求解内部机制复杂的问题；

2、网络能通过学习带正确答案的实例集自动提取“合理的”求解规则，即具有自学习能力；

3、网络具有一定的推广、概括能力。

6.3.2 BP 神经网络的问题

1、BP算法的学习速度很慢，其原因主要有：

(1) 由于BP算法本质上为梯度下降法，而它所要优化的目标函数又非常复杂，因此，必然会出现“锯齿形现象”，这使得BP算法低效；

(2) 存在麻痹现象，由于优化的目标函数很复杂，它必然会在神经元输出接近0或1的情况下，出现一些平坦区，在这些区域内，权值误差改变很小，使训练过程几乎停顿；

(3) 为了使网络执行BP算法，不能用传统的一维搜索法求每次迭代的步长，而必须把步长的更新规则预先赋予网络，这种方法将引起算法低效。

2、网络训练失败的可能性较大，其原因有：

(1) 从数学角度看，BP算法为一种局部搜索的优化方法，但它要解决的问题为求解复杂非线性函数的全局极值，因此，算法很有可能陷入局部极值，使训练失败；

(2) 网络的逼近、推广能力同学习样本的典型性密切相关，而从问题中选取典型样本实例组成训练集是一个很困难的问题。

3、难以解决应用问题的实例规模和网络规模间的矛盾。这涉及到网络容量的可能性与可行性的关系问题，即学习复杂性问题。

4、网络结构的选择尚无一种统一而完整的理论指导，一般只能由经验选定。为此，有人称神经网络的结构选择为一种艺术。而网络的结构直接影响网络的逼近能力及推广性质。因此，应用中如何选择合适的网络结构是一个重要的问题。

5、新加入的样本要影响已学习成功的网络，而且刻画每个输入样本的特征的数目也必须相同。

6、网络的预测能力（也称泛化能力、推广能力）与训练能力（也称逼近能力、学习能力）的矛盾。一般情况下，训练能力差时，预测能力也差，并且一定程度上，随训练能力地提高，预测能力也提高。但这种趋势有一个极限，当达到此极限时，随训练能力的提高，预测能力反而下降，即出现所谓“过拟合”现象。此时，网络学习了过多的样本细节，而不能反映样本内含的规律。

7、机器运行，无法解释步骤。

8、网络隐含节点数的确定至今没有任何理论上指导，尽管有一些经验公式，但仅供参考而已，具体问题还是要具体分析。

§7 总结与展望

7.1 总结

本文在研究四种非平稳时间序列建模与预测方法的基础上,对其中的两种算法提出改进:一是在小波分析时对逼近系数的重构信号使用更合理的 $ARIMA$ 法建模,通过差分使其平稳后,再建立 $ARMA$ 模型,而不是当作平稳序列直接建模,这样更符合信号的实际情况;二是对经过 $Kalman$ 滤波后分解出趋势项的使用 $ARMA$ 法建模,将 $Kalman$ 滤波过程与 $ARMA(p,q)$ 模型结合进行外推预测,克服了 $Kalman$ 预测不收敛的缺陷,并提高了预测精度。此外对非平稳时间序列的灰度组合模型和BP神经网络模型的建模与预测也进行了研究,四种算法都用 *Matlab* 编写了相应程序,以同样的上证综合指数序列的预测为例,各自得出了预测结果。

四种算法的比较如下:

1、误差均方根值

四种方法预测误差均方根值的比较如表7.1所示:

表7.1 四种方法预测误差均方根值的比较

预测方法	小波分解	$Kalman - ARMA$	灰度组合模型	BP神经网络
预测误差均方根值	95.6065	65.0260	172.4857	113.5959

对比可知,在误差均方根值方面基于小波分析的预测方法和 $Kalman - ARMA$ 法要优于灰度组合模型和BP神经网络模型法。

2、曲线拟合

四种方法预测结果与实际数据的比较图如图7.1~7.4所示:

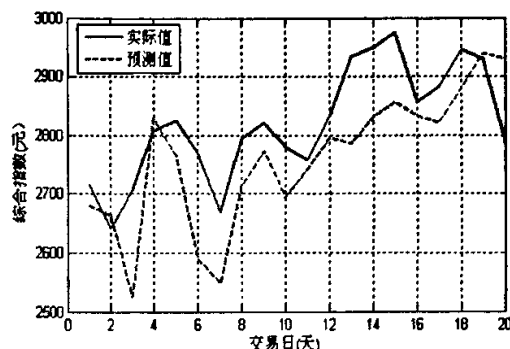


图7.1 小波分解预测法预测曲线拟合图

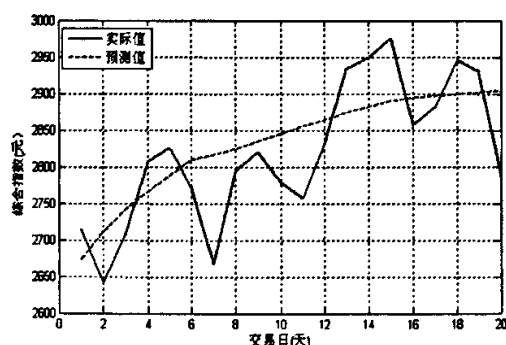


图7.2 Kalman-ARMA法预测曲线拟合图

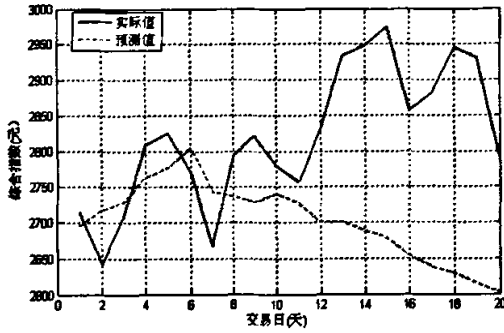


图7.3 灰度组合模型预测曲线拟合图

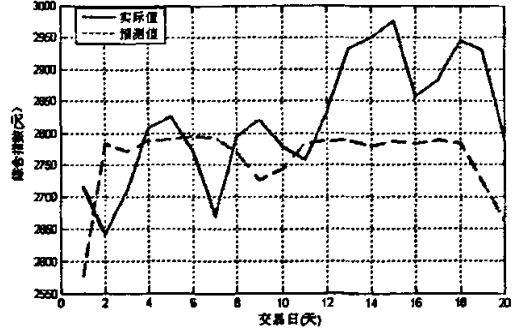


图7.4 BP神经网络预测曲线拟合图

对比可知,基于小波分析的预测方法和 $Kalman-ARMA$ 法的曲线拟合度要明显好于灰度组合模型和BP神经网络。在曲线波形方面基于小波分析预测方法的波形最接近于实际曲线波形。

3、计算复杂度

在计算复杂度方面,采用 *Matlab* 神经网络工具箱函数后,BP神经网络最易于实现;同样利用递推公式的灰度组合模型和 $Kalman-ARMA$ 法计算量相当;由于对每一个细节系数的重构信号都要分别建立 $ARMA$ 模型,对逼近系数的重构信号要建立 $ARIMA$ 模型,计算复杂度随着小波分解层数的增多而增大,因此四种方法中基于小波分析的预测方法计算最为复杂。

综上,基于小波分析的非平稳序列建模方法优点是预测精度高,曲线拟合好,缺点是计算复杂。 $Kalman-ARMA$ 预测法优点是预测精度高,能较准确反应预测趋势,缺点是总体趋势确定下,预测曲线仍有一定随机性。灰度预测模型的优点是计算简便,缺点是误差高,不能很好反应真实情况。BP神经网络模型的优点是模型建立简便,缺点是黑箱作业,无法监控预测过程,预测结果具有随机性。

7.2 展望

时变AR模型是正在兴起的另一非平稳时间序列的建模方法,相对于 $ARIMA$ 模型它不需要经过差分平稳就可以直接建立模型,避免了差分对序列趋势项有用信息的丢失,减少了建模步骤,提高了模型的可信度(参见文献[36])。如果采用小波分析法时结合时变AR模型则可以更进一步提高预测的精度和可信度。另外, $Kalman$ 滤波过程中,状态转移矩阵 F 的正确选择,不仅可以有效的分解变量而且可以更好的利用 $Kalman$ 递推过程,这些都是我需要进一步研究的方向。

总之,非平稳随机信号的分析方法正处在迅速发展。虽然它对科学技术发展的深远意义和影响还难以预料,但它的发展则是肯定的。本文只是对非平稳时间序列的建模和预测方法作了粗线条的勾画,扼要地介绍了相关的基础知识,分析了四种主要的预测方法并对其中两种提出了自己的改进意见,以期引起更多的研究者对这一领域的关注。在此应当

指出的是，不同的分析方法都各有其优点，同时也都存在局限性，因此在应用时，应根据具体情况选择合适的分析方法。而且，两种或多种分析方法的有机结合可能更有意义。

参考文献

- [1] 胡广书. 数字信号处理—理论、算法与实现 (第二版). 北京: 清华大学出版社, 2003, 8.
- [2] 姚天任, 孙洪. 现代数字信号处理. 华中科技大学出版社, 1999, 11.
- [3] 飞思科技产品研发中心. MATLAB 7 辅助信号处理技术与应用. 电子工业出版社, 2005, 3.
- [4] (美)George E.P.Box, (英)Gwilym M.Jenkins, (美)George C.Reinsel 著, 顾岚主译. 时间序列分析: 预测与控制. 中国统计出版社. 1997, 9.
- [5] (美)布洛克威尔 等著. 时间序列的理论与方法 (第二版). 田铮 译. 高等教育出版社; 施普林格出版社. 2001, 7.
- [6] Stephane G.Mallat . A theory for multiresolution signal decomposition :The wavelet representation [J]. IEEE Tran. on Pattern Analysis and Machine Intelligence ,1989 ,11 (7) :674 - 693.
- [7] Ingrid Daubechies. Orthonormal bases of compactly supported wavelets [J]. Communication on Pure and Applied Mathematics ,1988 ,41(11) :909 - 996.
- [8] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems [J]. Transactions of the ASME—Journal of Basic Engineering, 82 (Series D): 35-45.
- [9] 葛哲学, 陈仲生. Matlab 时频分析技术及其应用. 人民邮电出版社, 2006, 1.
- [10] 李国勇. 智能控制及其 MATLAB 实现. 电子工业出版社, 2005, 5.
- [11] 易丹辉编著. 数据分析与EViews应用. 北京: 中国统计出版社, 2002, 10.
- [12] 邓自立. 最优估计理论及其应用: 建模、滤波、信息融和估计. 哈尔滨工业大学出版社, 2005, 6.
- [13] 王丽娜, 肖冬荣. 基于ARMA模型的经济非平稳时间序列的预测分析. 武汉理工大学学报, 2004, 1.
- [14] 徐科, 徐金梧, 班晓娟. 基于小波分解的某些非平稳时间序列预测方法. 电子学报, 2001, 4.
- [15] 常晓丽. 基于Matlab的BP神经网络设计. 机械工程与自动化, 2006, 4.
- [16] 韩路跃, 杜行检. 基于MATLAB的时间序列建模与预测. 计算机仿真, 2005, 4.
- [17] 王文学, 王强, 孙萍. 卡尔曼滤波在机器人足球比赛系统中的应用. 机器人, 2006, 7.
- [18] Garcia-Trevino, E. S., Alarcon-Aquino, V. Single-step prediction of chaotic time series using wavelet-networks. 2006 Electronics, Robotics and Automotive Mechanics Conference 6 pp. 2006.
- [19] Cheng-Jian Lin. Wavelet neural networks with a hybrid learning approach. Journal of

- Information Science and Engineering vol.22, no.6: 1367-87 Nov. 2006.
- [20] Murat, Y.S., Ceylan, H. Use of artificial neural networks for transport energy demand modeling. Energy Policy vol.34, no.17: 3165-72 Nov. 2006.
- [21] 叶阿忠, 李子奈. 我国通货膨胀的混合回归和时间序列模型. 系统工程理论与实践, 2000, 9.
- [22] 梁来存, 皮友静. 我国保费收入的ARIMA模型与预测. 统计与决策, 2006, 7.
- [23] 邓聚龙. 灰色系统基本方法[M]. 武汉: 华中理工大学出版社, 1996.
- [24] 徐东升. 基于GM(1, 1)-ARMA 的软土地基沉降预测. 中国科技论文在线, 2006, 7.
- [25] 向书坚. 20世纪90年代时间序列预测领域主要研究动态. 中南财经大学学报, 2001, 2.
- [26] 王波, 张凤玲. 神经网络与时间序列模型在股票预测中的比较. 武汉理工大学学报, 2005, 12.
- [27] 盛艳波. 基于BP神经网络和ARIMA组合模型预测浙江省人均国内生产总值. 商场现代化, 2006, 8.
- [28] de Souza, L. V., Pozo, A. T. R., Neto, A. C.. Using correlation to improve boosting technique: an application for time series forecasting. 2006 18th IEEE International Conference on Tools with Artificial Intelligence 4 pp. 2006.
- [29] Kizilkaya, A., Kayran, A.H.. ARMA model parameter estimation based on the equivalent MA approach. Digital Signal Processing vol.16, no.6: 670-81 Nov. 2006.
- [30] Ortigueira, M. D., Serralheiro, A. J.. A new least-squares approach to differintegration modeling. Signal Processing vol.86, no.10: 2582-91 Oct. 2006.
- [31] Karrasch, C., Fiedeldej, M., Biermann, J., et al.. Forecasting of electrical load. ETZ vol.127, no.2: 22-6 Feb. 2006.
- [32] Wagenmakers, E.-J., Grunwald, P., Steyvers, M.. Accumulative prediction error and the selection of time series models. Journal of Mathematical Psychology vol.50, no.2: 149-66 April 2006.
- [33] 高紫光, 路磊. 非平稳时间序列的状态空间建模与预测. 系统工程, 1998, 5.
- [34] 陈颖, 李在铭. 一种改进的分段平稳随机过程的参数估计方法. 电子与信息学报, 2003, 6.
- [35] Al-Smadi, A.. The estimation of the order of an ARMA process using third-order statistics. International Journal of Systems Science vol.36, no.15: 975-82 15 Dec. 2005.
- [36] 张海勇, 马孝江, 盖强. 一种新的时变参数AR模型分析方法. 大连理工大学学报, 2002, 3.

致 谢

衷心的感谢我的导师吕振肃教授。本文的研究工作是在吕老师的悉心指导下完成的，从论文的选题、研究计划的制定、技术路线的选择到论文的撰写，各个方面都离不开吕老师的热情教导和帮助。攻读硕士学位期间，吕老师渊博的学术知识，深厚的科研功底，严谨的治学态度和敏锐的学术洞察力，都给我留下了难以磨灭的印象，也为我今后的工作和学习树立了优秀的榜样。因此，值此论文顺利完成之际，我要向尊敬的吕老师表达深深的敬意和无以言表的感谢。

感谢和我一起学习、生活了三年的同学们，你们给予了我无私的帮助和鼓励，这美好的三年将是我今后岁月的珍藏。

感谢我的父母和我的妻子，你们真挚的爱给了我不断前进的勇气和动力，我会继续努力，不辜负你们的期望。

作者：[管鹏](#)
学位授予单位：[兰州大学](#)

相似文献(10条)

1. 期刊论文 [马社祥, 刘贵忠, 曾召华, MA She-xiang, LIU Gui-zhong, ZENG Zhao-hua](#) [基于小波分析的非平稳时间序列分析与预测 -系统工程学报2000, 51 \(4\)](#)

提出了基于小波变换的非平稳时间序列分析预测方法. 通过小波分解, 将原时间序列依尺度分解成不同层次, 使趋势项、周期项和随机项分离, 对每一层进行分析与预测, 最后再合成得到原时间序列的预测值. 实例验证该方法是可行的.

2. 期刊论文 [周志杰, 胡昌华, 韩晓霞, ZHOU Zhi-jie, HU Chang-hua, HAN Xiao-xia](#) [基于非平稳时间序列的陀螺漂移性能建模与预测方法研究 -电光与控制2005, 12 \(3\)](#)

研究了利用Kalman滤波和小波分析两种方法对陀螺漂移非平稳时间序列建模预测的基本思想和具体算法. 实验证明, 这两种方法较传统的差分法优越, 可以很好地完成非平稳时间序列的建模预测, 并且通过比较分析了它们各自的优缺点, 为选用合理的建模方法提供了依据.

3. 学位论文 [吴学森](#) [小波分析理论在疾病流行动态定量分析中的应用研究](#) 2004

该研究利用山东省平邑县1985年~2002年HFRS月发病率资料, 在遵循流行病学原理的基础上, 将小波分析理论和数理统计方法有机结合, 定量研究了HFRS的流行动态变换趋势:①选择满足正则条件的小波基函数, 对时间序列数据作多层小波分解;并运用数学中的奇异性理论和小波分析的模极大值原理识别疾病流行的暴发时点;结合HFRS监测点的鼠情资料分析暴发原因;②根据资料的特点选择恰当的小波基函数和分解层数, 分离资料中的趋势项、周期项和随机项, 以获得疾病流行的短期波动、周期性变化趋势和长期趋势;③在②中小波分解的基础上, 应用时间序列模型或曲线拟合等方法分别对平邑县、泰安市和日照市HFRS资料分离后的各项作出预测, 再把各项的预测值相加得出各地HFRS月发病率的预测值. 综上所述, 小波分析理论对HFRS流行动态趋势中暴发时点的识别是灵敏和准确的;识别出的周期变化特征能较好地刻画HFRS动态变化趋势;小波分析与其它建模方法相结合能有效地提高非平稳时间序列预测精度, 是处理非平稳时间序列预测问题的有效方法, 为解决非平稳时间序列预测问题提供了新的思路. 该研究首次将小波分析理论应用到疾病流行动态定量分析中, 旨在为探讨影响疾病流行动态变化的危险因素及其强度和研究疾病未来流行变化趋势的预测, 探索一种科学的研究方法. 提出出的HFRS流行特征为进一步探讨疾病流行状态发生变化的原因和制定防治措施提供了理论依据. 该研究不仅能丰富疾病预测的理论和方法, 也拓展了小波分析的应用领域, 具有理论意义和实用价值.

4. 期刊论文 [翁宜慧, 季小俊, 孙志挥](#) [基于小波分析的预测算法在智能商务中的应用 -计算机工程2003, 29 \(10\)](#)
针对智能商务系统数据挖掘的特性, 应用小波分析理论, 提出了基于小波分析的预测算法, 并将其运用于“天商-2000”智能商务系统中.

5. 期刊论文 [吴学森, 王洁贞, 刘云霞, 张娜](#) [基于小波分析的肾综合征出血热发病率预测方法 -中国卫生统计](#) 2005, 22 (1)

目的建立季节性趋势时间序列小波预测模型的方法, 提高肾综合征出血热发病率的预测精度. 方法对原始序列进行多层小波分解, 提取趋势项、周期项和随机项, 对它们分别进行预测, 将各项的预测值合并作为原序列的预测值. 结果小波分解后建模的预测精度为84. 8%, 而ARIMA建模的预测精度为66. 1%. 结论用小波预测模型对HFRS作短、中期预测是有效、可行的.

6. 学位论文 [季小俊](#) [基于小波分析的数据挖掘方法及其在智能商务中的应用](#) 2002

小波分析方法作为信号分析的一种方法, 由于其时间和频率的分析特性, 在时间序列模式挖掘等领域有着广泛的应用前景. 而对于数据挖掘系统来说, 小波分析可以直接对数据进行预处理如去噪、频率突变点判断等. 该文在进行小波分析理论的研究前提下, 着重分析了一维连续小波变换和一维离散小波变换, 并结合数据挖掘系统要求, 提出了数据挖掘预处理的算法框架. 并针对非平稳时间序列提出了一下几种小波预测算法模型: (1) 基于小波分析的AR预测算法模型; (2) 基于小波分析的灰色系统预测模型; (3) 基于小波分析的灰色系统预测改进模型; 结合天商-2000智能商务系统, 检验了文中提出的数据预处理算法, 以及相应的预测算法. 实验证明文中提出的算法是有效可行的.

7. 期刊论文 [李盘荣, 马纲, 侯宁](#) [运用小波分析对非平稳时间序列的灰色预测 -机电技术2007, 30 \(2\)](#)

作者提出一种基于小波分析理论的灰色预测方法, 以此方法将非平稳时间序列分解到不同尺度上以减少原始序列的随机性, 然后用传统的灰色预测模型对重构后的时间序列分别进行预测, 最终得到预测值. 利用该方法对旋转机械零件的裂纹扩展进行了动态预测, 取得了良好的效果.

8. 学位论文 [柯熙政](#) [混沌、分形及小波分析的若干应用](#) 1999

该文根据混沌、分形以及小波分析的基本理论, 研究了测量数据中的非线性特性, 得出了一些普遍性的结论. 主要工作包括:1、分析了原子钟噪声中的混沌现象, 表明环境条件对原子钟工作状态有很大影响. 2、用复小波分析研究了地球极周期变化, 所得结果与前人的研究成果可互相印证. 3、提出了非平稳时间序列建模的新方法—分域递推模型, 从数学上进行了严格证明, 并将该方法应用与太阳黑子数和极移的分析. 4、用小波矩量法研究了粗糙面的电磁散射问题, 对各种周期粗糙面和随机粗糙面进行了分析, 并讨论了高阶矩特性. 5、用现代谱估计方法研究了粗糙面的电磁散射特性, 结果表明:这种方法具有很高的实用价值. 6、研究了分形粗糙面的后向增强效应, 对有关实验现象进行了分析研究, 所得结果与实验现象相当吻合.

9. 期刊论文 [徐科, 徐金梧, 班晓娟](#) [基于小波分解的某些非平稳时间序列预测方法 -电子学报2001, 29 \(4\)](#)

提出一种时间序列预测方法, 称为小波预测方法. 通过小波分解可以将某些非平稳时间序列分解成多层近似意义上的平稳时间序列, 然后采用自回归模型对分解后的时间序列进行预测, 从而得到原始时间序列的预测值. 对年平均太阳黑子数的预测结果表明, 该方法比传统的时间序列预测方法和神经网络预测方法的预测精度高, 可以很好地应用于某些非平稳时间序列的预测中.

10. 期刊论文 [王晓](#) [小波神经网络及其应用 -科技创新导报2009, "" \(27\)](#)

本文将某时间段的传染病数据视为非平稳的时间序列, 提出了一种小波多分辨分析和小波神经网络相结合的传染病预测方法, 一定程度上克服了现有的传染病预测方法过度依赖数学模型的局限性. 对美国疾病控制中心(CDC)2003-2004年的疑似流感(ILI)%数据仿真实验表明了该方法对于非平稳的时间序列预测的有效性.

本文链接: http://d.g.wanfangdata.com.cn/Thesis_Y1089226.aspx

授权使用: 西安电子科技大学(xadzkj), 授权号: 59b62d57-6d72-4396-9518-9deb00afb08b

下载时间: 2010年9月8日