

赵辰阳,徐 明. 基于 FIG-SVM 的农产品价格趋势预测[J]. 江苏农业科学,2014,42(5):385-388.

基于 FIG-SVM 的农产品价格趋势预测

赵辰阳,徐 明

(江苏师范大学教育研究院,江苏徐州 221000)

摘要:针对传统预测方法对长时预测效果不理想且无法做到趋势预测的问题,提出了一种基于 FIG-SVM 的农产品市场价格预测方法。利用粒化计算的思想将农产品市场价格序列提升到上层粒度序列,通过 SVM 对新形成的粒化序列进行预测,形成预测区间,最后根据实测值与预测值的对比结果判断预测方法的优劣。以全国大豆市场平均价格作为试验数据,结果表明,该方法预测精度较高,能够应用于实际场景。

关键词:FIG-SVM;模糊信息粒化;支持向量机;农产品价格;趋势预测

中图分类号:F304.3 **文献标志码:**A **文章编号:**1002-1302(2014)05-0385-03

农产品市场是我国市场经济的重要组成部分,其主要表现形式与核心要素是农产品价格^[1]。通过农产品的历史价格数据预测未来农产品价格,对指导农产品生产以及调整农业产业结构具有重要意义。目前,关于农产品价格的预测研究多以提高预测准确性为主,主要方式为线性建模与非线性建模 2 类^[2-3]。线性建模预测的常用方法包括移动平均法、指数平滑法、趋势外推法、自回归分析法等^[4-6];非线性建模预测的常用方法主要包括神经网络预测、支持向量机预测、混沌时间序列预测等^[7-9]。然而,在实际情况中对未来某个时间段市场价格的变化范围预测比对实际价格值预测更具实际意义。传统的预测模型如神经网络、SVM 等方法对长时预测效果并不理想,且无法做到对趋势范围的预测。针对这个问题,本研究提出了一种基于 FIG-SVM 的农产品价格趋势预测方法。利用粒化计算的思想将农产品市场价格序列提升到上层粒度序列,通过 SVM 对新形成的粒化序列进行预测,形成预测区间,最后根据实测值与预测值的对比结果判断预测方法的优劣。结果表明,该方法预测精度较高,具有较好的鲁棒性和可靠性。

1 模糊信息粒化

1.1 信息粒化

信息粒化是粒化计算和词语计算的主要方面,研究信息粒化的形成、表示、粗细、语义解释等方面,是根据对象的不可区分性、功能相近性、相似性、函数性等原则将对象划分为不同集合的方法,1 个集合为 1 个信息粒^[10]。在信息粒化中,非模糊的信息粒化方法在许多领域起重要的作用,常见的方法包括区间信息粒化、相空间信息粒化、基于信息密度的信息粒化。但在更多情况中,非模糊的信息粒不能明确地反映所描述事物的特性,需要采用模糊信息粒化方法进行模糊处理。

1.2 模糊信息粒化方法

模糊信息粒就是以模糊集形式表示的信息粒。用模糊集

方法对农产品价格进行模糊粒化主要分为 2 个步骤:划分窗口和模糊化。划分窗口的主要目的是将数据集分成若干子序列,每个子序列作为一个独立的操作窗口;模糊化是将产生的每个窗口进行模糊化,生成时序模糊信息粒,此时生成的新序列称为粒化序列。

常用的模糊粒子包括三角型、梯型、高斯型等,本研究采用梯型模糊粒子^[11],其隶属函数可以表示为:

$$A(x, a, m, n, b) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{m-a} & x \in (a, m] \\ 1 & x \in (m, n] \\ \frac{b-x}{b-n} & x \in (n, b] \\ 0 & x > b \end{cases} \quad (1)$$

其中: m, n 为梯形模糊粒子的核, a 为梯形模糊粒子的下边界, b 为梯形模糊粒子的上边界,且 $0 < a < m < n < b$ 。

2 FIG-SVM 预测模型

SVM 的基本思想是将低维不可划分的数据提升到高维空间进行处理,即通过内积函数定义的非线性映射将低维输入空间提升到一个高维空间,在高维空间中寻找输入变量和输出变量之间的非线性关系。具体步骤为:给定一个训练样本集 $S = \{(x_i, y_i), x_i \in R^d, y_i \in R, i = 1, 2, \dots, l\}$,其中 l 为训练样本的个数, d 为样本数据的属性维数。SVM 的目标是寻找最优回归函数: $y = w^T x + b$,其中 w 为权重向量, x 是样本输入数据, b 为偏移量。这个问题可以转化为如下二次规划问题:

$$\begin{aligned} \min & \frac{1}{2} w^T w + c \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{s. t. } & \begin{cases} y_i - w x_i - b \leq \varepsilon + \xi_i \\ w x_i + b - y_i \leq \varepsilon + \xi_i^* \end{cases} \end{aligned} \quad (2)$$

其中: $i = 1, 2, \dots, n$; ε 为不敏感损失函数的误差; ξ_i, ξ_i^* 表示松弛因子,且 $\xi_i, \xi_i^* \geq 0$; c 为惩罚因子,表示对超出 ε 的惩罚度。公式(2)可进一步表示为:

$$\begin{aligned} \min & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle - \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) + \\ & \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) \end{aligned}$$

收稿日期:2013-09-13

基金项目:江苏省徐州市科技计划(编号:XM12B085)。

作者简介:赵辰阳(1988—),男,江苏徐州人,硕士研究生,从事时间序列分析与预测方面的研究工作。E-mail: dmshxzmc@163.com。

$$\text{s. t. } \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i, \alpha_i^* \leq c \end{cases} \quad (3)$$

相应的回归函数为:

$$f(x) = w^T x + b = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle x_i, x_j \rangle + b^* \quad (4)$$

其中: α_i, α_i^* 为拉格朗日乘子, $\alpha_i - \alpha_i^* \neq 0$ 对应的样本是支持向量。以上的求解方式是针对线性情况。对于非线性情况, 在高维特征空间中, 线性问题中的内积运算可以用核函数来代替, 常用的核函数有多项式核函数、径向基核函数、多层感知器核函数等。研究成果与具体实践表明 RBF 核函数对非线性、高维数据具有较好的应用效果, 其数学表达式为:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \text{ 其中 } \sigma \text{ 为核宽度。核函数的参}$$

数 g 和 c 的选取是影响 SVM 预测效果的主要因素^[12], 本研究采用交叉验证的方式来选择最佳的 g 和 c , FIG-SVM 的运算流程如图 1 所示。

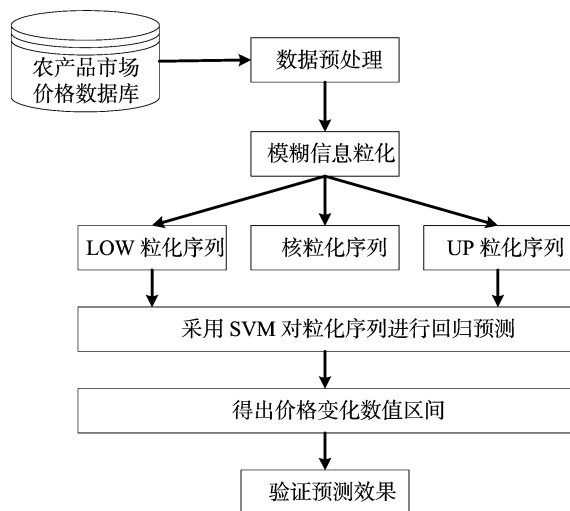


图1 FIG-SVM 模型的运算流程

其中模糊信息粒化采用 Pedrycz 提出的模糊粒化方法^[12], 其用于建立的模糊粒子的基本思想是^[12]: (1) 模糊粒子能够合理地代表原始数据, 即最大化 $\sum_{i=1}^k A(x_i)$; (2) 模糊粒子要有一定的特殊性, 也就是最小化模糊集 A 的支撑, 即最小化 $|b - a|$, 其中 b 和 a 分别是模糊集 A 的支撑边界。

通过以上基本思想得到如下优化函数:

$$Q = \frac{\sum_{i=1}^k A(x_i)}{|b - a|} \quad (5)$$

粒化的目标是使函数 Q 取得最大值。对于给定的时间序列粒化窗口 $X = \{x_1, x_2, \dots, x_N\}$, 则 Pedrycz 的模糊粒化算法 (FIG_P 算法) 具体步骤为: (1) 确定梯形模糊集的核 m, n 。将 $X = \{x_1, x_2, \dots, x_N\}$ 按由小到大的顺序排序, 不妨设排序后的时间序列仍为 $X = \{x_1, x_2, \dots, x_N\}$, 当 N 是偶数时, $m = X(N/2), n = X[(N+2)/2]$; 当 N 是奇数时, $m - n = X[(N+1)/2]$ 。 (2) 确定梯形模糊集支撑下界 a 。采用数据遍历的方法计算 Q , 使 Q 取最大值的数据值即为支撑下界 a , 即使函数 $Q(a) = \sum_{i=1}^k A(x_i) / (m - a)$ 取最大值的 a 为 A 的支撑下

界, 其中 $A(x_i)$ 为 x_i 的隶属度函数。 (3) 确定梯形模糊集支撑上界 b 。采用数据遍历的方法计算 Q , 使 Q 取最大值的数据值即为支撑上界 b , 即使函数 $Q(b) = \sum_{i=1}^k A(x_i) / (b - n)$ 取

最大值的 b 为 A 的支撑下界, 其中 $A(x_i)$ 为 x_i 的隶属度函数。

经过上述步骤可以确定梯形模糊信息粒子的 4 个参数为 a, m, n, b 。由每个序列窗口粒化后的支撑下界 a 形成的新的粒化序列记为 UP 序列, 由每个序列窗口粒化后的支撑下界 b 形成的新粒化序列记为 LOW 序列。采用 SVM 对 LOW 序列和 UP 序列进行预测可以最终得到未来市场价格的变化区间。

3 案例分析

3.1 数据

本研究采用的数据来自 2009 年 3 月 4 日至 2013 年 7 月 31 日共计 1 610 d 的全国大豆市场价格数据, 结果如图 2 所示。选用前 1 400 个数据作为训练集, 后 210 个数据为测试集, 进行多步预测。

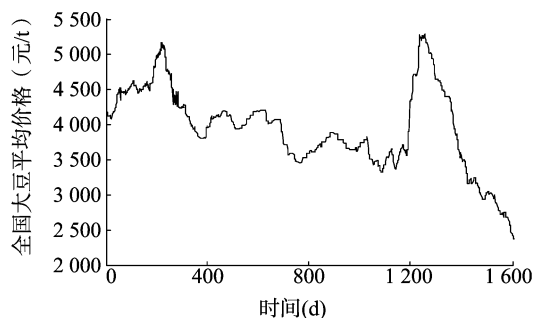


图2 全国大豆平均价格数据图

3.2 对原始数据进行模糊信息粒化

采用 Pedrycz 提出的模糊粒化方法对原始数据进行模糊信息粒化, 模糊粒子采用梯形模糊粒子, 窗口大小为 7, 即预测 1 周内的价格变化趋势与区间范围。为了能够更好地看出粒化后的覆盖度, 将粒化后的序列按照窗口 W 进行扩展, 形成高粒度序列在原始序列中的投影序列, 设粒化后的序列为 $Y = \{y_i\} = \{y_1, y_2, \dots, y_H\}$, 则投影序列 $Y_p = \{y_{ij}\} = \{y_{11}, y_{12}, \dots, y_{1w}, y_{21}, y_{22}, \dots, y_{2w}, y_{31}, \dots, y_{Hw}\}$, 其中 $y_i = y_{ij}, i = 1, 2, \dots, H, j = 1, 2, \dots, w$ 。模糊信息粒化效果如图 3 所示, 所显示的数据范围为 1 ~ 250。由图 3 可以看出, 粒化后的 UP 序列和 LOW 序列的投影序列包含原始数据。

3.3 对粒化序列进行回归预测

为了对 UP 序列和 LOW 序列实现更好的预测效果, 首先对粒化序列进行归一化处理, 将其归一到区间 $[0, 1]$ 之间。采用交叉验证的方式选取参数 g 和 c , 然后利用 SVM 对 UP 序列和 LOW 序列进行回归预测, 对测试集进行模糊信息粒化, 形成测试集的真实粒化序列, 预测值与实际值的对比效果如图 4、图 5 所示。

为了更好地评价预测效果, 分别绘制 UP 序列、LOW 序列预测值与实际值的相对误差比较图, 结果如图 6、图 7 所示。

由图 4 至图 7 可以看出, 预测的整体趋势满足实际情况, 且预测值与实际值相对误差较小, UP 序列的相对误差在

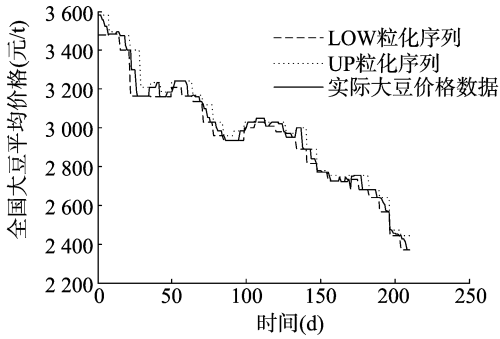


图3 模糊粒化投影效果

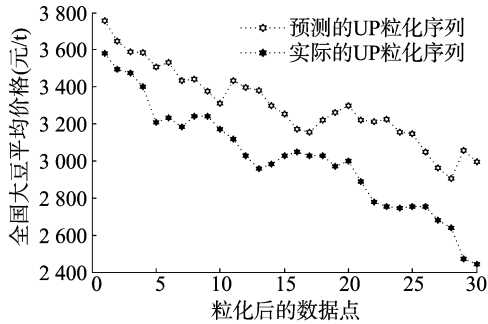


图4 UP 预测值与实际值对比结果

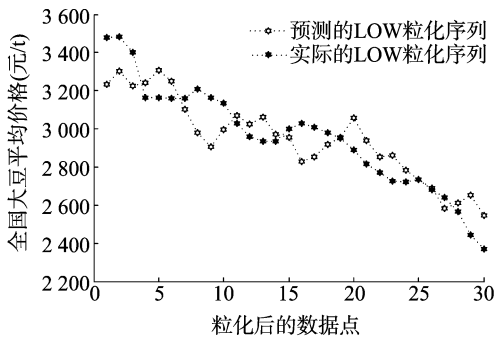


图5 LOW 预测值与实际值对比结果

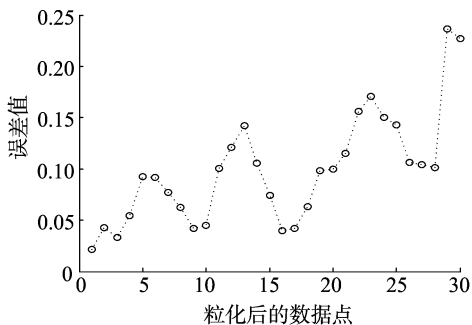


图6 UP 预测值与实际值相对误差情况

0.25 以内,LOW 序列的相对误差在 0.1 以内,表明基于模糊信息粒化的 SVM 趋势预测具有较好的试验效果,能够应用到实际应用中。

3.4 预测效果验证

为了进一步说明 FIG-SVM 的预测效果,检验 2013 年 1 月 2—8 日的大豆价格是否在上述预测范围内,如表 1 所示。

由表 1 可以看出,2013 年 1 月 2—8 日(1 周内)全国大豆

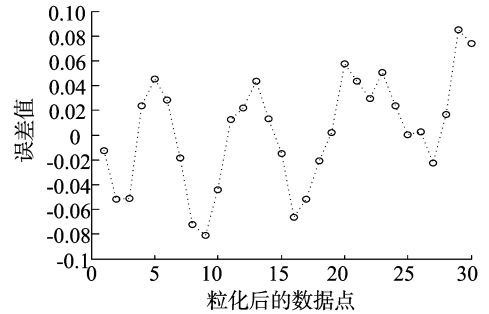


图7 LOW 预测值与实际值相对误差情况

表 1 全国大豆价格变化预测验证结果(2013 年 1 月)

日期	当日实际价格(元/t)
2 日	3 580
3 日	3 580
4 日	3 555
5 日	3 555
6 日	3 525
7 日	3 525
8 日	3 480

注:2013 年 1 月 2—8 日全国大豆价格预测范围为 3 435.46 ~ 3 656.55 元/t。

价格均在 FIG-SVM 预测的数值区间内,表明利用 FIG-SVM 对农产品价格趋势进行预测是可行而准确的。

4 结论

针对农产品价格每日不断变化难以确定其规律的问题,本研究提出基于模糊信息粒化的 SVM 回归预测方法,可以实现对农产品价格的数值空间预测;全国大豆市场价格真实数据的实际应用结果表明,该方法切实有效,能够为农业相关人员提供理论依据与价值参考。

参考文献:

- [1] 孙 倩,穆月英. 蔬菜价格变动、影响因素及价格预测——以北京市批发市场为例[J]. 中国蔬菜,2011(9):9-14.
- [2] 董晓霞,李干琼,刘自杰. 农产品市场价格短期预测方法的选择及应用——以鲜奶零售价格为例[J]. 山东农业科学,2010(1):109-113.
- [3] 任伟宏. 农产品市场价格预测方法探析[J]. 中国农学通报,2011,27(26):209-212.
- [4] 韩 雯. ARIMA 模型在贵州省农产品价格预测中的应用——以辣椒为例[J]. 安徽农业科学,2011,39(21):13226-13227,13229.
- [5] 刘 峰,王儒敬,李传席. ARIMA 模型在农产品价格预测中的应用[J]. 计算机工程与应用,2009,45(25):238-239,248.
- [6] 刘轶芳,迟国泰,余方平,等. 基于 GARCH-EWMA 的期货价格预测模型[J]. 哈尔滨工业大学学报,2006,38(9):1572-1575.
- [7] 康 璐,陈 欢,张蕾妮. 基于 GA-BP 神经网络模型的期货价格预测与分析[J]. 财经界,2011(9):108-109.
- [8] 罗长寿. 基于神经网络与遗传算法的蔬菜市场价格预测方法研究[J]. 科技通报,2011,27(6):881-885,894.
- [9] 孟 军,孙 超. 基于支持向量机的大豆产量预测研究[J]. 数学的实践与认识,2011,41(18):144-149.

冀 宏,赵黎明. 区域循环农业发展的技术范式升级策略——基于江苏省蒋巷村循环农业发展实证研究[J]. 江苏农业科学,2014,42(5):388-391.

区域循环农业发展的技术范式升级策略 ——基于江苏省蒋巷村循环农业发展实证研究

冀 宏^{1,2}, 赵黎明¹

(1. 天津大学管理与经济学部, 天津 300072; 2. 常熟理工学院, 江苏常熟 215500)

摘要:基于技术范式的视角,对江苏省蒋巷村循环农业发展进行实证研究。通过对蒋巷村三产循环的区域循环经济发展模式归纳与分析,阐明由于注重培育循环农业发展主体的生态认知和构建三产循环的经济和技术共同体,形成了跨越不同产业领域的生态链,实现了产业经济主体之间物质和能量的循环利用。明确指出经济生产者的空间集聚——区域型循环园区建设是促进循环农业技术范式形成的有效载体。在此基础上,进一步提出了区域性循环农业技术范式升级的具体措施:强化循环意识,提升循环农业技术范式认知;聚合多元产业集群,构建循环农业技术范式共同体;集成技术支撑体系,构筑循环农业技术范式升级轨道。

关键词:循环农业;技术范式;三产循环;共同体;区域循环园区;策略

中图分类号: F327 **文献标志码:** A **文章编号:** 1002-1302(2014)05-0388-04

将技术哲学领域的技术范式理论融入主流经济学,建立经济增长的生态价值内生化逻辑框架,构建了研究循环农业发展的技术范式新体系^[1-2]。在技术范式的视角下,不同的范式决定了不同的经济发展方式^[3]。人类社会在经济发展过程中经历了3种不同的技术范式:原始技术范式、现代技术范式、循环技术范式^[4]。从技术和经济的角度看,农业循环经济实际上是传统技术范式的变革和升级。

过去由于忽视对循环农业经济范式的整体考虑与设计,在循环农业的具体实践中,从源头上就已经欠缺可持续发展的条件和动力机制,从而导致我国循环农业发展存在诸多制约因素^[5]。循环农业技术范式使我们能够在主客体意义上对农业循环经济发展进行总括性的动力学把握,便于理解和掌握其内在规律及演进方式。本研究基于技术范式的视角,以江苏省蒋巷村循环农业发展实践为例,进行实证研究,提出区域循环农业技术范式升级的措施与建议,不仅对循环农业技术范式研究体系具有实践支撑作用,而且对其他地区发展农村循环经济具有现实的借鉴意义。

1 蒋巷村循环农业发展实践

蒋巷村位于江南腹地江苏省常熟市,全村186户,村户籍

人口800人,村辖面积3 km²。全村经济总产值超过12亿元,利税实现9 000万元,人均国内生产总值持续超2万美元,村民人均年收入2万元,先后被授予“全国文明村”“国家级农村现代化建设示范村”“江苏省百佳生态村”“循环经济示范村”等光荣称号,是江苏省政府批准的200个环境整治试点村之一,实现了农村经济与生态文明的协调发展。

蒋巷村循环农业发展确定的基本思路是“以发展循环经济为核心,融生态农业、生态工业及生态村建设为一体的环境可持续发展的现代化之路”,并依据农村产业结构演变规律,结合其优美的江南水乡文化特色,确定了“农业起家、工业发家、旅游旺家”的战略定位,实现了通过农业起家夯实经济发展基础,依靠工业发家强村富民,借助旅游旺家实现全面生态化发展,形成了以农业发展为根本、以工业反哺为驱动、以生态旅游为特色、以高效节能技术为手段、以循环经济为目标的发展基本框架^[6-7](图1),初步创建了一个经济发达地区循环农业建设新模式。

2 蒋巷村发展循环农业的经验

蒋巷村的循环农业发展之路没有陷入片面追求景观环境美化和形象工程的束缚,从建设伊始就在发展规划中把握了循环农业的经济发展本质和生态协调内涵,始终坚持发展与环保并重,将农业、工业、第三产业建设融为一体统筹规划,构建循环经济发展的共同体互为支撑,协同发展,立足农村自然条件和经济文化背景,注重对村域范围内空间发展的合理引导和土地利用的有效控制,始终保持中国特色的农村生态系统空间结构的可持续发展,并结合当地实际情况,因地制宜地

收稿日期:2013-09-04

基金项目:江苏省苏州市社会发展研究项目(编号:SSD201002)。

作者简介:冀 宏(1969—),男,河北保定人,博士研究生,研究员,主要从事技术经济及管理领域农业经济方向研究。Tel:(0512) 52251185;E-mail:jihong@cslg.cn。

[10] Li Y, Yu F S. Optimized fuzzy information granulation based machine learning classification [C]//2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery, 2010:259-263.

[11] Qiao L, Liang S, Cui F. Research on support vector machine prediction on surrounding rock deformation based on fuzzy information

granulation [C]//International Conference on Consumer Electronics, Communications and Networks, 2011:3466-3469.

[12] Cheng X, Guo P. Short-term wind speed prediction based on support vector machine of fuzzy information granulation [C]//25th Chinese Control and Decision Conference, 2013:1918-1923.