*Operations Research and Fuzziolgy* 运筹与模糊学, 2011, 1, 29-33
http://dx.doi.org/10.12677/orf.2011.12006 Published Online November 2011 (http://www.hanspub.org/journal/orf/)

# Grey LS-SVM Forecasting with Parameter Optimized by Genetic Algorithm

**Deqiang Zhou**
Department School of Information and Mathematic, Yangtze University, Jingzhou
Email: zdqmfk@yahoo.com.cn
Received: Oct. 5th, 2011; revised: Nov. 6th, 2011; accepted: Nov. 8th, 2011.

**Abstract:** This paper utilized the accumulation generation operation of grey prediction to produce accumulated data, and accumulated data were used to construct grey LS-SVM. At the same time the parameters for LS-SVM were pretreated through genetic algorithms to get the optimum parameter values, then the optimized LS-SVM based on genetic algorithms was used to small samples forecasting. A typical example was taken to be analyzed and compared with GM (1, 1) and LS-SVM method. The result shows that the method forecast effect is better, and the prediction model has better generalization ability.

**Keywords:** Grey Least Square Support Vector Machines; GM (1, 1) Model; Genetic Algorithms; Parameter Selection; Small Samples Forecasting

# 基于遗传算法优选参数的灰色 LS-SVM 预测

**周德强**
长江大学信息与数学学院，荆州
Email: zdqmfk@yahoo.com.cn
收稿日期：2011 年 10 月 5 日；修回日期：2011 年 11 月 6 日；录用日期：2011 年 11 月 8 日

**摘　要：**利用灰色预测方法中累加生成运算形成累加数据，将累加数据作为训练样本构造灰色 LS-SVM，并利用遗传算法对灰色 LS-SVM 自身的参数进行优选，然后将基于遗传算法优选参数的灰色 LS-SVM 用于小样本预测。选取了典型例子进行验证，并与传统 GM(1, 1)和 LS-SVM 方法进行对比。结果表明本文所提出的方法预测效果良好，且预测模型具有更好的泛化能力。

**关键词：**灰色 LS-SVM；GM(1, 1)模型；遗传算法；参数优选；小样本预测

## 1. 引言

目前，对时间序列的预测有很多种方法，从传统的 ARMA 模型到智能化的神经网络模型，这些方法的核心在于建立一个好的预测模型[1]。但是由于实际应用中时间序列具有不规则、混沌等非线性特性，加之样本的有限性，很难对系统建立理想的预测模型，因此，往往难以获得精确的预测结果。

将不同机理的预测方法相结合，为提高预测精度提供了新的途径。文献[2]提出一种串联灰色神经网络，将神经网络和灰色预测方法结合，互相取长补短，避免了 GM(1, 1)模型存在的理论缺陷[2]。但是，该方法仍然遗留了常规神经网络存在的缺陷，比如学习方法采用经验风险最小化(ERM)[3]准则，即最小化训练样本点误差，因而不可避免地出现过拟合现象，影响了模型的泛化能力。文献[1]提出灰色支持向量机预测方法，将支持向量机(SVM)和灰色预测方法结合，发挥了灰色预测方法中“累加生成”的优点，弱化了原始序列中随机扰动因素的影响，增强了数据的规律性，不仅避免了灰色预测方法及模型存在的理论缺陷，而且克服了神经网络模型中存在的局部极小问题。然而该方法也仍然遗留了常规支持向量机存在的缺陷，比如，模型中核函数以及模型参数的选择凭经验选取，

容易导致参数选择不准确而使最后的预测精度低于目标精度[4,5]。另一方面 SVM 在处理二次寻优问题上比较耗时，收敛速度较慢[6]。

Suykens 在 SVM 的优化函数中引入方差项，并将 SVM 中的不等式约束条件改为等式约束，提出了一种以二次等式约束条件为基础的改进型向量机，即最小二乘支持向量机(LS-SVM)[7]，LS-SVM 通过方差项的引入，将经典 SVM 优化函数的不等式约束改成了只有等式约束，这样 LS-SVM 的求解问题从标准 SVM 的二次函数寻优问题转换为线性方程求解问题，而且不再需要指定逼近精度 $\varepsilon$，大大简化了问题的复杂性，求解速度相对加快[5,6]。

通过分析灰色预测与 LS-SVM 预测各自的优缺点及互补性，本文将两种小样本预测技术结合，构造灰色 LS-SVM，并利用遗传算法[5,8,9]对灰色 LS-SVM 自身的参数进行自动搜索和确定，再利用经寻优预处理的灰色 LS-SVM 模型进行预测。计算实例表明本文的方法是可行的且有效的，比传统方法预测精度高，可用于小样本预测。

## 2. GM(1, 1)和 LS-SVM 的比较

### 2.1. GM(1, 1)预测方法

设 $X^{(0)}=\left(x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n)\right)$ 表示原始数据序列。对原始序列作一阶累加生成得到新序列

$$X^{(1)}=\left(x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n)\right),$$

其中 $x^{(1)}(k)=\sum_{j=1}^{k}x^{(0)}(j),(k=1,2,\cdots,n)$。对新序列的预测值 $\hat{X}^{(1)}$ 利用逆累加生成得到原始序列 $X^{(0)}$ 的灰色预测值，表示为

$$\hat{X}^{(0)}(k+1)=X^{(1)}(k+1)-X^{(1)}(k)=\left(1-e^{a}\right)\left[X^{(0)}(1)-\frac{u}{a}\right]e^{-ak},k=1,2\cdots.$$

这里，规定 $\hat{X}^{(0)}(1)=X^{(0)}(1)$，其中 $a,u$ 为 GM(1, 1)模型参数。

GM(1, 1)预测方法所需样本数较少，该模型通过原始数据的累加可以消除一定的随机性，但它并不是一种万能方法，对于有些数据模型的解并不合理，理论上要求原始数据累加生成近似指数型序列，对于具有波动性变化的非线性序列，其预测精度往往不能令人满意，适用范围受到很大限制[1,2,10]。

### 2.2. LS-SVM 预测方法

对于训练样本 $\left\{(x_k,y_k)\right\}_{k=1}^{n},x_k\in \mathrm{R}^n,y_k\in \mathrm{R}$，用非线性映射 $\phi(x)$ 将样本从原空间映射到一个维数为 $k$ 的高维特征空间 $Z$ 中，LS-SVM 的目标是在该空间中构造最优的线性回归函数

$$f(x)=w^T\phi(x)+b, \tag{1}$$

式中 $w\in \mathrm{R}^k$ 为权向量，$b\in \mathrm{R}$ 为偏移量。

根据 SRM 原则，LS-SVM 算法表述为[7]

$$\min_{w,b,\xi}\frac{1}{2}\|w\|^2+\frac{\lambda}{2}\sum_{k=1}^{n}\xi_i^2 \quad \text{s.t.}\quad y_k=w^T\phi(x_k)+b+\xi_k,k=1,2,\cdots n \tag{2}$$

其中 $\xi_k$ 为误差项，$\lambda$ 是一个调节因子，当 $\lambda$ 为无穷大时，所得的解为最小二乘解。实际计算中，不需要知道非线性变换 $\phi(x)$ 的具体形式，可用核函数 $K(x_i,x_j)$ 来实现算法的线性化。

为求解回归函数，引入如下拉格朗日函数

$$L(w,b,\xi,\alpha)=\frac{1}{2}\|a\|^2+\frac{\lambda}{2}\sum_{k=1}^{n}\xi_i^2-\sum_{k=1}^{n}\alpha_k\left[w^T\varphi(x_k)+b+\xi_k-y_k\right], \tag{3}$$

式中 $\alpha_k$ 为拉格朗日乘子。根据 KKT 优化条件有

$$\frac{\partial L}{\partial a}=0,\frac{\partial L}{\partial b}=0,\frac{\partial L}{\partial \xi_k}=0,\frac{\partial L}{\partial \alpha}=0, \tag{4}$$

得到如下等式约束条件

$$\begin{cases} w=\sum_{k=1}^{n}\alpha_k\varphi(x_k), \\ \sum_{k=1}^{n}\alpha_k=0, \\ \alpha_k=\lambda\xi_k, \\ w^T\varphi(x_k)+b+\xi_k-y_k=0 \end{cases} \tag{5}$$

消去 $w$ 和 $\xi$，问题归结为求解如下线性方程组

$$\begin{bmatrix} 0 & \boldsymbol{e}^T \\ \boldsymbol{e} & \Omega+\dfrac{I}{\lambda} \end{bmatrix}_{n\times n}\begin{bmatrix} b \\ \alpha \end{bmatrix}=\begin{bmatrix} 0 \\ y \end{bmatrix} \tag{6}$$

式中 $\boldsymbol{e}$ 为元素为 1 的 $(n-1)\times 1$ 向量，$I$ 为 $(n-1)\times(n-1)$ 的单位阵，$\alpha=[\alpha_2,\alpha_3,\cdots,\alpha_n]^T$，$y=[y_2,y_3,\cdots,y_n]^T$，$\Omega=\left(k\left(x_i,x_j\right)\right)_{(n-1)\times(n-1)}$。

求方程组(6)的最小二乘解为

$$\begin{bmatrix} b^* \\ \alpha^* \end{bmatrix}=\left(\begin{bmatrix} 0 & \boldsymbol{e}^T \\ \boldsymbol{e} & \Omega+\dfrac{I}{\lambda} \end{bmatrix}_{n\times n}^T \begin{bmatrix} 0 & \boldsymbol{e}^T \\ \boldsymbol{e} & \Omega+\dfrac{I}{\lambda} \end{bmatrix}_{n\times n}\right)^{-1} \begin{bmatrix} 0 & \boldsymbol{e}^T \\ \boldsymbol{e} & \Omega+\dfrac{I}{\lambda} \end{bmatrix}_{n\times n}^T \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{7}$$

得 LS-SVM 线性回归函数为

$$f(x)=\sum_{k=1}^{n}\alpha_k^* K(x_k,x)+b^* \tag{8}$$

支持向量机严格的数学基础使其在理论上有较大优势，但与其理论研究相比，应用研究则相对滞后[1]。同时，LS-SVM 预测精度在很大程度上依赖于训练集的选择[11]。核函数以及模型参数的选取至今没有一定的理论做指导[1]。研究者往往凭经验和有限的实验给定一组参数[4]。

### 2.3. GM(1, 1)与 LS-SVM 预测方法的互补性

灰色预测适合处理“小样本”、“贫信息”的数据，LS-SVM 适合处理“小样本”、“非线性”的数据。两种方法都适合解决小样本的预测问题，这是共性，然而预测方法却存在较大差异，但分析各自特点，存在以下互补性：

1) 灰色预测方法是一种不严格的系统方法，它避开系统结构分析环节，直接通过对原始数据的累加构建指数增长模型，寻找系统的整体规律，但不适合逼近复杂的非线性函数。LS-SVM 则具有逼近任意函数的能力。

2) LS-SVM 虽然对样本数量依赖性弱，但 LS-SVM 预测精度在很大程度上依赖于训练集的质量，而灰色预测方法适合处理“贫信息”的数据，对数据集要求不高。

3) LS-SVM 的预测性能对于参数的选择比较敏感[4]，利用灰色预测方法中“累加生成”的优点，削弱原始数据序列中随机扰动因素的影响，使离乱的原始数据中蕴涵的规律充分显露出来，增强数据的规律性，得到便于 LS-SVM 学习的具有单调增长规律的新序列[1]，也同时便于用优化方法寻找核函数的参数。灰色预测方法中的参数 $a,u$ 对 GM(1, 1)预测精度也有较大影响[1]，利用 LS-SVM 算法建立预测模型，不再求解 GM(1, 1)模型中的参数，可避免 GM(1, 1)模型求解参数的理论缺陷。

因此，在灰色预测和 LS-SVM 预测都适合解决小样本问题的共性下，将两种方法结合，取长补短，形成性能更好的预测方法具有可行性。

## 3. 灰色 LS-SVM 预测方法

### 3.1. 灰色 LS-SVM 模型

根据上述对 GM(1, 1)和 LS-SVM 特点的分析，本文提出仅保留灰色预测方法中“累加生成”和“累减还原”运算，利用 LS-SVM 算法建立灰色 LS-SVM 预测模型，不再求解 GM(1, 1)模型中的参数。具体算法设计如下，

1) 对原始数据序列 $X^{(0)}=\left(x_1^{(0)},x_2^{(0)},\cdots,x_n^{(0)}\right)$ 作累加生成：$x_k^{(1)}=\sum_{j=1}^{k}x_j^{(0)},(k=1,2,3,\cdots,n)$，得到生成序列 $X^{(1)}=\left(x_1^{(1)},x_2^{(1)},\cdots,x_n^{(1)}\right)$，并构成训练样本 $\left\{\left(k,x_k^{(1)}\right),k=1,2,\cdots,n\right\}$。

2) 选择核函数 $K(i,j)$，利用训练样本构造灰色 LS-SVM 模型，表示为

$$\begin{aligned} &\min_{a,b,\xi}\frac{1}{2}\|w\|^2+\frac{\lambda}{2}\sum_{k=1}^{n}\xi_i^2 \\ &\text{s.t.}\quad x_k^{(1)}=w^T\varphi(k)+b+\xi_k,k=1,2,\cdots n \end{aligned} \tag{9}$$

3) 利用 LS-SVM 算法求解模型(9)，构造回归函数，$\hat{x}_k^{(1)}=\sum_{j=1}^{n}\alpha_j^* K(j,k)+b^*$。

4) 利用累加序列的预测值 $\hat{x}^{(1)}$，进行“累减还原”得到原始序列的灰色预测值 $\hat{x}^{(0)}$。

### 3.2. 基于遗传算法优化求解灰色 LS-SVM 参数

目前研究最多的核函数主要有三类:多项式核函数、径向基核函数(RBF)、Sigmoid 核函数。此外，核函数自身的参数和调节参数 $\lambda$ 的选取对应用结果也有较大影响，为达到最佳预测效果，以 RBF 核函数

$K\left(x_i,x_j\right)=\exp\left(-\left\|x_i-x_j\right\|_2^2/\sigma^2\right)$ 为例说明利用遗传算法[11-13]对灰色 LS-SVM 的参数 $\sigma$ 和 $\lambda$ 进行优选。

以误差平方和最小化为目标，即

$$\min_{\sigma,\lambda}\sum_{k=2}^{n}\left(\hat{x}^{(0)}(k)-x^{(0)}(k)\right)^2 \tag{10}$$

其中 $\hat{x}^{(0)}(k)$ 为灰色 LS-SVM 模型预测值，设计求解最优 $\sigma$ 和 $\lambda$ 的遗传算法如下[5,12]：

1) 编码表示。将参数 $\sigma\in(0,+\infty)$ 和 $\lambda\in(0,+\infty)$ 表示为 $n$ 位二进制串( $n$ 的大小可根据 $\sigma$ 和 $\lambda$ 的精度来确定)。

2) 构造初始群体。取种群大小 $N$，在 $(0,+\infty)$ 上随机生成 $N$ 对 $\sigma$ 和 $\lambda$ (染色体)作为初始种群 $P(0)=\left\{\left[\sigma(1,0),\lambda(1,0)\right],\left[\sigma(2,0),\lambda(2,0)\right],\cdots,\left[\sigma(N,0),\lambda(N,0)\right]\right\}$

3) 对每个染色体解码，采用如下适应度函数 $\mathrm{Fit}\left(\sigma(i,k),\lambda(i,k)\right)=1\Big/\sum_{j=1}^{n}\left(\hat{x}^{(0)}_{\sigma,\lambda(i,k)}(j)-x^{(0)}(j)\right)^2$ 计算群体 $P(k)$ 中每个染色体 $\sigma(i,k)$，$\lambda(i,k)$ (其中 $k\le K$ 表示代数，$K$ 为最大进化代数)的适应度。其中 $\hat{x}^{(0)}_{\sigma,\lambda(i,k)}(j)$ 表示由参数 $\sigma(i,k)$，$\lambda(i,k)$ 得到的 GM(1, 1)模型的预测值。

4) 对种群进行遗传操作。计算每代中各个个体的生存概率 $p_i^{(k)}=\mathrm{Fit}\left(\sigma(i,k),\lambda(i,k)\right)\Big/\sum_{j=1}^{M}\mathrm{Fit}\left(\sigma(j,k),\lambda(j,k)\right)$ $M$ 表示第 $k$ 代种群规模。设计一个随机选择策略(如“赌盘选择”)，使每个个体被选择进行繁殖的概率为 $p_i^{(k)}$，将繁殖生成的个体组成父代 $P(k+1)$。

以概率 $P_c$ 对新一代个体进行交叉换位，产生新的个体，再对新个体以概率 $P_m$ 进行变异操作。

5) 反复执行步骤(3)~(4)，直至得到满意解或已达到了预设的最大代数为止。

## 4. 基于遗传算法优选参数的灰色 LS-SVM 的预测方法

通过上述分析，提出基于遗传算法优选参数的灰色 LS-SVM 的预测方法，具体算法设计如下，

1) 对原始数据序列 $X^{\pi(0)}=\left(x_1^{(0)},x_2^{(0)},\cdots,x_n^{(0)}\right)$ 作累加生成：$x_k^{(1)}=\sum_{j=1}^{k}x_j^{(0)},(k=1,2,3,\cdots,n)$，得到生成序列 $X^{(1)}=\left(x_1^{(1)},x_2^{(1)},\cdots,x_n^{(1)}\right)$，并构成训练样本 $\left\{\left(k,x_k^{(1)}\right),k=1,2,\cdots,n\right\}$。

2) 选择核函数 $K(i,j)$(比如 RBF 核)，利用训练样本构造灰色 LS-SVM 模型(9)。

3) 利用遗传算法求解式(10)，对灰色 LS-SVM 回归模型(9)中的参数进行寻优预处理。

4) 对经过优化参数的灰色 LS-SVM，依据 LS-SVM 算法求解，构造回归函数 $\hat{x}_k^{(1)}=\sum_{j=1}^{n}\alpha_j^*K(j,k)+b^*$。

5) 利用累加序列的预测值 $\hat{x}^{(1)}$，进行“累减还原”得到原始序列的灰色预测值 $\hat{x}^{(0)}$。

## 5. 实例分析

利用文[13]中给出的 1997 年到 2003 年火灾伤人率数据进行预测，1997~2000 年火灾伤人率数据作为样本集，2001~2003 年火灾伤人率数据作为测试集，见表 1。

分别用本文方法(简称累加 LS-SVM)，原始序列 LS-SVM 模型(简称原始 LS-SVM)进行建模。为具有可比性，用遗传算法优化累加 LS-SVM 模型和原始 LS-SVM 模型的参数时，控制参数设置为相同的，种群大小 $N=50$，二进制编码长度为 20，交叉概率为 0.95，变异概率为 0.08，最大进化代数 $K=250$。用遗传算法求得累加 LS-SVM 模型的参数为 $\sigma=12.028$，$\lambda=897.17$，原始 LS-SVM 模型的参数 $\sigma=21.019$，$\lambda=637.04$。利用平均绝对相对误差作为评价指标，与 GM(1, 1)模型的拟合结果进行比较，结果见表 2。

本文的方法在测试集上的平均相对误差为仅为 7.10%，比用 GM(1, 1)和 LS-SVM 方法所得的平均相

**Table 1. Data of fire injury rate of China from 1997-2003**
**表 1. 我国 1997~2003 年火灾伤人率数据($10^{-6}$)**

| 年份 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 |
|---|---|---|---|---|---|---|---|
| 伤人率 | 4 | 3.9 | 3.7 | 3.5 | 2.96 | 2.66 | 2.38 |

**Table 2. Analysis and prediction of fire injury rate of China**
**表 2. 我国 1997~2003 年火灾伤人率预测分析**

| 平均绝对相对误差(%) | 样本集 | 测试集 |
|---|---|---|
| 累加 LS-SVM 模型 | 2.40 | 7.10 |
| 原始 LS-SVM 模型 | 0.15 | 21.67 |
| GM(1, 1)模型 | 4.58 | 12.20 |

对误差分别降低 5.10%和 14.57%，预测精度得到极大的改善。虽然原始序列适合 GM(1, 1)模型，GM(1, 1)模型能够描述数据内部变化的本质，但传统的 GM(1, 1)模型本身存在理论缺陷，使得模型并没有显示良好的推广能力，而本文的模型发挥了灰色预测方法中“累加生成”的优点，削弱了原始数据中的随机性，增强了规律性，相对直接利用 LS-SVM 模型显示更好的推广能力，同时避免了 GM(1, 1)方法本身存在的理论缺陷，预测结果显示模型发挥了 GM(1, 1)和 LS-SVM 两种小样本技术的优势。

## 6. 结论

综上所述，可得如下结论:

1) 将 GM(1, 1)和 LS-SVM 这两种“小样本”技术结合形成灰色 LS-SVM，可集成 2 种技术在处理小样本数据上的优点，使单一 LS-SVM 方法和 GM(1, 1)各自的不足得到互补。

2) 在灰色 LS-SVM 模型参数的选择上，利用遗传算法对模型自身的参数进行自动搜索和确定，可克服传统 LS-SVM 参数选择方法中存在的缺点，保证最终的预测模型具有更好的泛化能力。

3) 理论分析和实际应用表明本文提出的方法是可行的且有效的，可用于小样本预测。

## 参考文献 (References)


[1] 唐万梅. 基于灰色支持向量机的新型预测模型[J]. 系统工程学报, 2006, 21(4): 410-413.

[2] 张大海, 毕研秋, 毕研霞等. 基于串联灰色神经网络的电力负荷预测方法[J]. 系统工程理论与实践, 2004, 12: 128-132.

[3] V. N. Vapnik. The nature of statistical learning theory. Heidellberg: Springer Verlag, 1995.

[4] 吴景龙, 杨淑霞, 刘承水. 基于遗传算法优化参数的支持向量机短期负荷预测方法[J]. 中南大学学报(自然科学版), 2009, 40(1): 180-184.

[5] 周辉仁, 郑丕谔, 赵春秀. 基于遗传算法的LS-SVM参数优选及其在经济预测中的应用[J]. 计算机应用, 2007, 27(6): 1418-1419, 1429.

[6] 王宇红, 黄德先, 高东杰等. 基于LS-SVM的非线性预测控制技术[J]. 控制与决策, 2004, 19(4): 383-387.

[7] J. A. K. Suykens, J. Vandewalle. Least squares support vector machine classfiers. Neural Processing Letters, 1999, 9(3): 293-300.

[8] J. H. Holland. Genetic algorithms. Scientific American, 1992, 4: 44-50.

[9] P. J. Denning. Genetic algorithms. Journal of Parallel and Computing, 1992, 80(1): 354-360.

[10] 李伟, 韩力. 组合灰色预测模型在电力负荷预测中的应用[J]. 重庆大学学报, 2004, 27(1): 36-39.

[11] 席裕庚, 耿晓军, 陈虹. 预测控制性能研究的新进展[J]. 控制理论与应用, 2000, 17(4): 469-475.

[12] 何文章, 宋国乡. 基于遗传算法估计灰色模型中的参数[J]. 系统工程学报, 2005, 20(4): 432-436.

[13] 陈子锦, 王福亮, 陆守香. 灰色预测模型GM(1, 1)的适用性分析及在火灾风险预测中的应用[J]. 中国工程科学, 2007, 9(5): 91-94.


 