

文章编号: 1001-0920(2006)09-1054-05

## 一种模糊支持向量负荷预测法及其参数优化策略

蒋刚<sup>1</sup>, 肖建<sup>1</sup>, 宋昌林<sup>2</sup>, 郑永康<sup>1</sup>

(1. 西南交通大学 电气工程学院, 成都 610031; 2. 四川大学 电气工程学院, 成都 610065)

**摘要:** 分析了电力系统负荷预测目前采用方法的不足, 并根据电网负荷的特点提出一种基于模糊支持向量的核回归方法, 同时提出多参数同步优化策略, 增强了该方法的实用性和有效性。从理论上分析了小样本条件下, 该方法可以有效避免过学习的原因, 它不需设计网络结构, 降低了对实验人员经验的依赖程度。与神经网络法进行对比实验, 实验结果表明了该方法的优越性和适用性, 并具有较好的实用价值和应用前景。

**关键词:** 模糊支持向量; 核回归; 负荷预测; 多参数同步优化; 统计学习理论

中图分类号: TM 7; TP391

文献标识码: A

## A Kind of Peak Load Forecasting Method Based on Fuzzy Support Vector and It's Parameter Optimization

JIAN G Gang<sup>1</sup>, XIAO Jian<sup>1</sup>, SONG Chang-lin<sup>2</sup>, ZHENG Yong-kang<sup>1</sup>

(1. School of Electric Engineering, Southwest Jiaotong University, Chengdu 610031, China; 2. School of Electric Engineering, Sichuan University, Chengdu 610065, China. Correspondent: JIAN G Gang, E-mail: jgg4648@sohu.com)

**Abstract:** The shortcomings of current used methods in load forecasting of power systems. A fuzzy support vector kernel regression method is presented. A multi-parameters synchronous optimization is also presented to enhance the effectiveness of the method. Experiments based on power system data show that this kind of kernel regression method brings lower error, can avoid over fitting effectively, no need to design neural network structure, and use fewer time than artificial neural network. It seems to be the most promising one and has some more applied value.

**Key words:** Fuzzy support vector; Kernel function regression; Load forecasting; Multi-parameters synchronous optimization; Statistic learning theory

### 1 引言

电网负荷预测是电力系统的一项基本工作, 对开停机的宏观调控具有重要的指导作用。预测的精度直接影响到整个电力系统运行的经济效益, 根据历史负荷数据和相关条件预测电网的负荷是一个值得深入研究的课题。

电网负荷受多种因素影响, 关系错综复杂, 模糊, 非线性, 强扰动, 难于精确建模。自20世纪80年代以来, 国内外众多专家和学者对负荷预测和相关技术进行了大量的理论分析和实例应用研究。目前

采用的方法主要有: 时序分析法<sup>[1]</sup>、最小二乘法<sup>[2]</sup>、灰色理论法<sup>[3]</sup>和人工神经网络法, 其中又以神经网络应用最为广泛。神经网络以经验风险最小化(ERM)为优化目标, 训练的时候可以有极小的误差甚至零误差, 但对未经训练的新数据, 往往表现出较差的推广能力, 存在过学习的先天不足, 小样本情况下这个缺陷尤其突出。此外, 局部极值点、收敛速度慢、隐含层结构难于确定等缺点也较大地限制了其应用。网络结构的选择主要通过实验的方式反复凑, 这个过程通常要花费较长的时间, 在很大程度上依赖于实验人员的经验。

收稿日期: 2005-06-23; 修回日期: 2005-10-06

基金项目: 教育部博士点基金项目(20040613013)。

作者简介: 蒋刚(1978—), 男, 四川广安人, 博士生, 从事统计学习理论与智能控制的研究; 肖建(1950—), 男, 湖南衡阳人, 教授, 博士生导师, 从事计算机控制、智能控制等研究。

针对神经网络的缺陷, 研究人员提出了改进方法<sup>[4,5]</sup>, 并取得了较好的效果, 从一定程度上提高了预测精度, 产生了较好的经济效益, 但在算法原理上仍然缺乏实质性的突破, 因此难以从根本上改善神经网络自身的不足

针对负荷的模糊、非线性、强扰动等特点, 本文提出了一种模糊支持向量的核回归方法, 与 3 层神经网络进行对比实验, 结果表明该方法具有较强的鲁棒性和推广能力, 它不需要专门设计网络结构, 降低了对实验人员经验的依赖程度, 花费少量的训练时间即可达到某一精度指标, 对于多参数模型优化时间过长、不利于工程实际应用的问题, 本文提出了一种多参数同步优化策略, 成功地将优化时间缩减至原优化时间的 0.04%~ 0.06%, 使该方法的实用性和有效性得到大幅增强

## 2 模糊支持向量核回归方法

模糊支持向量核回归方法(F-SV KR)建立在统计学习理论(SLT)基础之上, 核函数为支持向量机(SVM)提供了重要的构成模块<sup>[6,7]</sup>.

给定样本集  $S = \{(x_i, y_i), i = 1, 2, \dots, n\}$ , 如果其输出是模糊的, 那么传统的核函数方法将不能对它求解, 因为无法直接求解它对应的最优分类超平面. 为此采取以下策略: 首先将原回归问题转化为有限维的模糊分类问题, 得到模糊样本集

$$S_F = \{(x_i, y_i \pm \delta_i, s_i); z_j\},$$

$$s.t. i = 1, 2, \dots, n, j = 1, 2, \dots, m, m \geq n. \quad (1)$$

其中:  $(x_i, y_i \pm \delta_i)$  是模糊输入样本,  $\delta_i$  是正实数,  $\pm \delta_i$  是  $y_i$  的左右展宽,  $s_i$  是该样本对应的模糊隶属函数,  $z_j = \pm 1$  是第  $j$  类的输出. 对多类分类问题, 目前采用的策略主要有<sup>[8]</sup>: One-against-one, One-against-all, Directed Acyclic Graph (DAG).

通过映射  $\Phi$  将原空间  $L$  映射到特征空间  $H$ , 对应使用核函数  $K$ , 令第  $k$  类样本中心为  $\bar{x}_k$ , 则该类样本中样本点  $x_{kt}$  的广义模糊隶属距离定义为以下形式:

$$d_k(t) = \sqrt{K(\bar{x}_k, \bar{x}_k) + K(\bar{x}_k, x_{kt}) + K(x_{kt}, x_{kt})}. \quad (2)$$

定义样本点  $x_{kt}$  的模糊隶属函数为

$$s_k(t) = d_k(t) / \max d_k \quad (3)$$

通过映射  $\Phi$  将原空间  $L$  映射到特征空间  $H$ , 令  $y_F = y \pm \delta$ , 它将导致权向量  $w$  和阈值  $b$  的模糊, 即  $w_F = w \pm \delta_w, b_F = b \pm \delta_b, \pm \delta_w$  和  $\pm \delta_b$  分别是  $w$  和  $b$  的左右展宽. 由此构建广义模糊最优分类超平面方程

$$w_F \Phi(x) + sy_F + b_F = 0 \quad (4)$$

在非线性且不可分的情况下, 引入松弛变量  $\xi$ , 得到约束表达式为

$$[z_j (w \Phi(x) + sy_F + b) - 1] + \xi = 0,$$

$$s.t. j = 1, 2, \dots, m. \quad (5)$$

原二次规划问题转化为如下形式的模糊决策机会约束规划(FCCP)<sup>[9]</sup>:

$$\min \{ \Phi(w, s, C, \xi) =$$

$$\left[ \frac{1}{2} (w_F^2 + s^2 + C \sum_{j=1}^m \xi_j) \right],$$

$$s.t. P \left[ \frac{1}{2} (w_F^2 + s^2 + C \sum_{j=1}^m \xi_j) \leq \eta \right] \geq \pi,$$

$P \{ 1 - z [w_f \Phi(x) + sy_F + b_F] \leq \gamma \} \geq \tau$   
 其中:  $C = f[d(j)]$  为惩罚因子, 定义为广义模糊隶属距离  $d$  的函数, 对不可分情况下的错分样本进行控制, 实现错分样本的比例与算法复杂度之间的折衷处理,  $f$  可设置为线性函数  $C = cd(j)$  形式;  $\xi = 0$  为松弛变量;  $\eta = 0$  和  $\gamma = 0$  是设定的阈值;  $\pi = 0, \tau = 0$  是一定的置信水平;  $P\{\cdot\}$  表征事件  $\{\cdot\}$  对应的概率测度

引入 Lagrange 乘子  $\alpha, \beta$ , 可得

$$\min [L(w, b, \xi)] =$$

$$\min \left\{ \frac{1}{2} w_F^2 + C \sum_{i=1}^n s_i \xi_i - \sum_{i=1}^n \alpha \xi_i - \sum_{i=1}^n \beta_i \{ \epsilon + \xi_i + 1 - z_i [w_F \mathcal{Q}(x_i) + s(i)y_F(i) + b_F(i)] \} \right\} \quad (6)$$

根据鞍点定理, 用  $L$  分别对  $w, b, \xi$  求偏导, 令其为零, 并将结果代入式(6), 得

$$\min \left\{ \frac{1}{2} \sum_{i,j=1}^n \left( \sum_{h=1}^k z_{ih} \alpha_{ih} - \alpha_{ih} \right) \left( \sum_{h=1}^k z_{jh} \alpha_{jh} - \alpha_{jh} \right) \mathcal{Q}(x_i) \mathcal{Q}(x_j) - \sum_{i,j=1}^n \left[ z_j \left( \sum_{h=1}^k (\alpha_{ih} - \alpha_{jh}) \mathcal{Q}(x_i) \mathcal{Q}(x_j) \right) + b_F - 2 \right] - \sum_{i=1}^n \sum_{h=1}^k [(\alpha_{ih} - C s_{ih} + \beta_{ih}) \xi_{ih}] \right\}$$

约束条件为

$$\alpha_{svm} = \sum_{i=1}^n \sum_{h=1}^k z_{ih} \alpha_{ih}$$

可得决策函数

$$f(x) =$$

$$\arg \max_{svm} \left[ \sum_{h=1}^k z_h \alpha_h - \alpha_{svm} \mathcal{Q}(x) \mathcal{Q}(x_i) + b \right] =$$

$$\arg \max_{svm} \left[ \sum_{h=1}^k z_h \alpha_h - \alpha_{svm} \right] K(x, x_i) + b$$

其中  $K$  为核函数, 常用的核函数有: 线性核  $K(x, x_i) = xx_i$ , 多层感知器核  $K(x, x_i) = \tanh(v xx_i + c)$ , 多项式核  $K(x, x_i) = (xx_i + 1)^2$ , 径向基核  $K(x, x_i) = \exp(-((x - x_i)/\sigma)^2)$  等

### 3 多参数同步优化策略

对多参数模型进行优化时, 最典型的问题是优化时间过长, 不利于实际工程应用 为此, 提出一种多参数同步优化策略, 实验结果验证了该策略的有效性

给定参数序列

$$\text{para} = \{p_i, i = 1, 2, \dots, n\},$$

各参数对应的取值范围为

$$\text{bound} = \{[p_{ia}, p_{ib}], i = 1, 2, \dots, n\},$$

优化时的搜索步长为

$$\text{step} = \{P_{is}, i = 1, 2, \dots, n\},$$

则内循环体总共需要执行的次数为

$$N_1 = \prod_{i=1}^n \left( \left\lfloor \frac{p_{ib} - p_{ia}}{p_{is}} \right\rfloor + 1 \right), \quad (7)$$

其中  $\lfloor \cdot \rfloor$  表示对  $\cdot$  做向下取整运算

可见, 即使内循环体执行一次只花费 1 s 时间, 每个参数只迭代 10 次, 但是在 6 个参数需要优化的情况下, 共需要花费大约 277. 78 h, 这在工程应用中是没有实用价值的

基于此, 考虑将参数序列 para 分别处理, 对参数  $p_k$  优化时, 固定其余  $(n - 1)$  个参数  $\{p_1, \dots, p_{k-1}, p_{k+1}, \dots, p_n\}$ . 定义误差指标为相对偏差百分比, 即

$$\text{Err} = \left[ \frac{|L_F - L_R|}{L_R} \right] \times 100\%, \quad (8)$$

其中:  $L_F$  为预测值,  $L_R$  为真实测量值 那么  $p_k$  的取值范围  $[p_{ka}, p_{kb}]$  将对应于图 1 所示的误差分布示意图

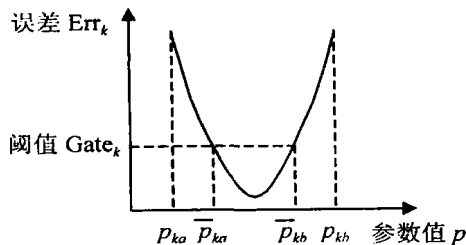


图 1 某取值范围对应的误差分布示意

设定某误差阈值  $\text{Gate}_k$ , 取低于该阈值对应的参数上下界  $[p_{ka}, p_{kb}]$ , 并设置新的搜索步长  $\bar{p}_{ks}$  每个参数均作此操作, 形成阈值截取面(如果每个参数的阈值设置为不同值, 则阈值截取面不是平面, 而是曲面), 如图 2 所示

采用上述方法, 一般经 4 ~ 5 次迭代即可收敛到较小的区间  $\text{bound}^* = \{[p_{ia}^*, p_{ib}^*], i = 1, 2, \dots, n\}$ , 这种同步优化思想的程序执行示意如图 3 所

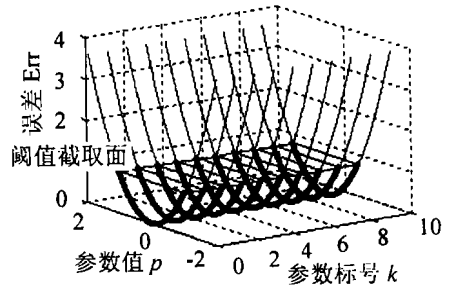


图 2 多参数形成的阈值截取面三维示意

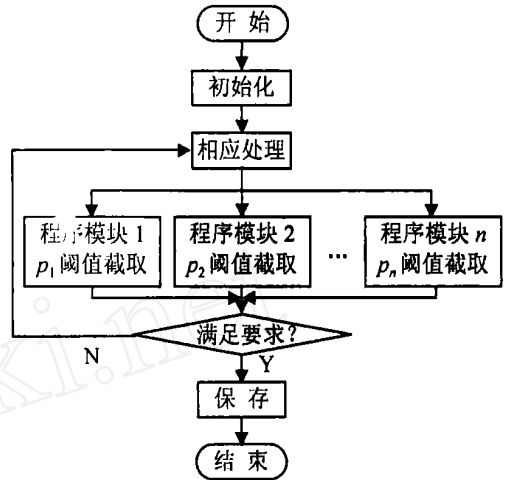


图 3 同步优化思想的程序执行示意

示

需要注意的是: 由于受当前计算机体系结构(冯·诺依曼机)的限制, 该程序在执行时是串行执行的, 不能真正实现“同步”, 如果采用生物计算机, 同步的思想可以实现, 但超出本文研究的范畴, 这里不做讨论 在实验中发现, 即使不能实现真正意义上的同步, 这种优化方法仍然体现出其优势, 可做进一步深入探讨

内循环体总共的执行次数为

$$N_2 = (4 \sim 5) \prod_{i=1}^n \left( \left\lfloor \frac{\bar{p}_{ib} - \bar{p}_{ia}}{\bar{p}_{is}} \right\rfloor + 1 \right). \quad (9)$$

这样效率将提高  $\eta = N_1/N_2$  倍, 加入附加程序的开销, 最终的优化时间将缩减至原优化时间的 0.04% ~ 0.06%, 从而使该方法在工程应用中的实用性和有效性得到大幅增强

## 4 预测实例

### 4.1 数据预处理

采用 F-SV KR 和 3 层人工神经网络法(ANN)对某地电网 2003 年 12 月 1 日至 2004 年 12 月 31 日的历史负荷数据进行对比实验 电网负荷受多种因素影响, 如节假日  $F$ , 日平均气温  $T$ , 日最高气温  $T_H$  和持续时间  $t_H$ , 日最低气温  $T_L$  和持续时间  $t_L$ , 空气湿度  $H$ , 当地居民生活水平提高  $L$ , 当地第一/第二

/ 第三产业的相对比例  $I_1, I_2, I_3$  及其增长因子  $G_1, G_2, G_3$ , 电能能源消耗中所占比重  $P$  等 由于是短期峰值负荷预测, 故只取  $F, T, T_H, T_L$ , 暂不考虑其他因素对负荷的影响

取预测日前 28 天的数据, 按是否为节假日分成两种数据类型, 2004 年 366 天, 形成 366 个可用数据集 每个非节假日数据集包含 22 ~ 24 个数据, 每个节假日数据集包含 4 ~ 6 个数据, 它们是预测日前 28 天对应数据类型的温度测量值和每天的最大负荷测量值, 按列向量排列成  $m \times n$  矩阵 对这种类型的时间序列作预测的时候, 认为越接近预测日的数据对预测的作用越大, 设置权重为

$$W(p) = a / ((b + 1)^p), \quad (10)$$

其中权重衰减因子  $a$  和  $b$  用于控制衰减速率,  $p$  是所用数据对应日期与预测日相隔的天数

考虑到温度、负荷等物理量的变化范围差异很大, 量纲也不同, 为了给高斯核函数标准差  $\sigma$  的选取提供参照系, 先对它们作标准化处理 然后根据式(1) ~ 式(3) 对数据作进一步处理, 将其转化为模糊样本集, 取得适当的模糊隶属函数

#### 4.2 参数分析

实验采用 3 层神经网络, 通过试凑法选择隐含层单元数; 对 F-SV KR 采用高斯核, 即

$$K(x, x_i) = \exp(-|x - x_i|^2 / \sigma^2),$$

核参数  $\sigma$ , 惩罚因子  $C = cd(j)$  的参数  $c, \epsilon$  管道半径对预测精度的影响非常大, 按上述介绍的多参数同步优化策略进行优化, 优化结果取值如下:  $\sigma = 4.2, c = 71.5, \epsilon = 0.0012$

1) 固定  $c$  和  $\epsilon, \sigma$  的取值区间为  $[0.1, 15]$ , Err 先减小后增大, 在 4.2 附近形成谷点

2) 固定  $\sigma$  和  $\epsilon, c$  的取值区间为  $[1, 150]$ , Err 先减小后增大, 在 71.5 附近形成谷点 表明  $c$  太小, 会产生过学习现象;  $c$  太大, 会产生欠学习现象, 对推广能力影响很大

3) 固定  $\sigma$  和  $c, \epsilon$  的取值区间为  $[0.0001, 1]$ , Err 先减小后增大, 在 0.0012 附近形成谷点 若  $\epsilon$  较小, 有较多支持向量数, 精度可提高, 但太小会产生过学习现象; 若  $\epsilon$  较大, 支持向量数较少, 精度降低, 但推广能力有所增强

4) 式(10) 中的权重因子  $a = 82, b = 237$ , 与其他参数相比,  $a$  和  $b$  具有较大的调节空间, 对预测精度的影响相对较小

#### 4.3 预测结果

取 366 个集中在 2 月份和 8 月份的数据为测试集, 其余 10 个月的数据作样本集进行训练 误差按式(8) 计算, 预测误差如表 1 所示

表 1 每日峰值负荷预测误差 %

日期	2月		8月	
	F-SV KR	ANN	F-SV KR	ANN
1	3.18	7.70	2.69	3.10
2	1.32	2.24	0.84	3.54
3	2.72	7.60	1.50	0.80
4	0.85	6.89	2.05	4.14
5	3.63	2.76	3.49	0.50
6	1.91	1.58	0.70	4.24
7	2.65	2.00	3.03	2.15
8	1.27	2.63	0.46	5.07
9	0.92	5.51	2.78	1.97
10	1.71	8.21	4.07	6.65
11	3.41	7.51	1.26	1.07
12	3.53	8.99	1.81	0.13
13	2.86	2.16	0.54	5.36
14	3.10	4.86	1.16	3.04
15	1.02	2.60	4.15	8.33
16	1.22	5.52	2.17	0.22
17	1.39	0.30	2.28	5.41
18	0.32	6.37	3.75	9.30
19	2.73	2.75	4.36	1.68
20	0.20	1.03	1.04	4.61
21	3.81	5.73	2.77	3.48
22	3.48	6.92	2.24	4.08
23	1.33	1.79	4.34	3.79
24	3.95	3.30	0.92	0.64
25	3.18	3.81	2.45	1.77
26	0.05	3.95	4.02	1.00
27	0.16	2.84	2.15	4.23
28	3.23	2.81	1.19	2.87
29	1.76	2.32	4.16	6.46
30	—	—	2.69	2.17
31	—	—	0.29	4.19
平均误差	2.10	4.23	2.30	3.41

### 5 结 论

针对电网负荷的特点提出了 F-SV KR; 对多参数优化耗时过长的的问题进行了分析, 提出了多参数同步优化策略, 成功地将优化时间缩短至原优化时间的 0.04% ~ 0.06%, 增强了该方法的实用性和有效性, 并对参数的选择和预测结果的内在关系作了详细的分析 与目前普遍采用的 ANN 进行短期峰值负荷预测的对比实验可以看出, F-SV KR 的预测精度高于 ANN, 表明对电网负荷这种模糊、非线性和强扰动系统, 本文提出的 F-SV KR 方法是一种行

## 之有效的方案

F-SV KR 以结构风险最小化SRM 为优化目标,同时控制经验风险和置信范围,即使在样本数较少的情况下,也表现出较强的推广能力;当样本较少时,以经验风险最小化ERM 为优化目标的神经网络法通常都会出现过学习现象,并且难以对其进行有效的控制 F-SV KR 求解的是一个二次规划问题,可以从理论上保证收敛到全局最优;而神经网络较容易收敛到局部极值点

采用核函数将非线性问题从原空间映射到高维特征空间,可以转化为线性问题进行求解,且可以越过映射的具体形式,有效避免维数灾难问题;神经网络通过启发式方法对特定问题进行求解,缺乏先验信息时,网络结构的确定难以有效解决,通常用实验的方式反复试凑,在很大程度上依赖于实验人员的经验 经过反复优化,神经网络虽然可以取得较高的精度,但是要达到同一精度指标,需要花费很多时间,而本文方法只需花费较少的训练时间即可做到

预测电网负荷时,F-SV KR 方法和多参数同步优化策略展示出良好的性能,具有较大的实用价值和良好的应用前景 分析其优势和可能存在的不足,并通过深入分析电网负荷的特点,构建性能良好的回归机是下一步要做的工作

## 参考文献(References)

- [1] Franklin P W. A Theoretical Study of the Three Phase Salient Pole Type Generator with Simultaneous AC and Bridge Rectified DC Output, Part I and Part II [J]. *IEEE Trans on Power Apparatus and Systems*, 1973, 92(2): 543-557.
- [2] Schiferl R F. Six Phase Synchronous Machine with AC and DC Stator Connections, Part I and Part II [J]. *IEEE Trans on Power Apparatus and Systems*, 1983, 102(8): 2685-2701.
- [3] Ma W M, Zhang G F, Liu D Z, et al. A Synchronous Machine with Simultaneous AC/DC Output [P]. China:

ZL 94107628.8, 1999-09-11.

- [4] 尤勇, 盛万兴, 王孙安. 基于人工免疫网络的短期负荷预测模型[J]. *中国电机工程学报*, 2003, 23(3): 26-30  
(You Y, Sheng W X, Wang S A. Short-term Load Forecasting Using Artificial Immune Network [J]. *Proc of the CSEE*, 2003, 23(3): 26-30)
- [5] 谢宏, 程浩忠, 张国立, 等. 基于粗糙集理论建立短期电力负荷神经网络预测模型[J]. *中国电机工程学报*, 2003, 23(11): 1-4  
(Xie H, Cheng H Z, Zhang G L, et al. Applying Rough Set Theory to Establish Artificial Neural Networks for Short Term Load Forecasting [J]. *Proc of the CSEE*, 2003, 23(11): 1-4)
- [6] Nello Cristianini, John S T. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods* [M]. Cambridge: Cambridge University Press, 2000
- [7] Vapnik V N. *Statistical Learning Theory* [M]. New York: Springer-Verlag, 2000
- [8] Hsu C W, Lin C J. A Comparison of Methods for Multi-class Support Vector Machines [J]. *IEEE Trans on Neural Networks*, 2002, 13(2): 415-425.
- [9] 赵晓煜, 汪定伟. 供应链中二级分销网络优化设计的模糊机会约束规划模型[J]. *控制理论与应用*, 2002, 19(2): 249-253  
(Zhao X Y, Wang D W. Fuzzy Chance Constrained Programming Model for Bi-level Distribution Network Design in the Supply Chain [J]. *Control Theory and Applications*, 2002, 19(2): 249-253)
- [10] 朱家元, 杨云, 张恒喜, 等. 支持向量机的多层动态自适应参数优化[J]. *控制与决策*, 2004, 19(2): 223-225  
(Zhu J Y, Yang Y, Zhang H X, et al. Multi-layer Adaptive Parameters Optimization Approach for Support Vector Machines [J]. *Control and Decision*, 2004, 19(2): 223-225)

(上接第1053页)

- [3] Bitmead R R, Gevers M, Wertz V. *Adaptive Optimal Control, the Thinking Man's GPC* [M]. Sydney: Prentice Hall, 1990
- [4] Alberto P, Ortega R. On Generalized Predictive Control: Two Alternative Formulations [J]. *Automatica*, 1989, 25(5): 753-755
- [5] Zhu K, Gorez R, Wertz V. Alternative Algorithms for Generalized Predictive Control [J]. *Systems and Control Letters*, 1990, 15(2): 169-173
- [6] Warwick K, Peterka V. Optimal Observer Solution for

Predictive and LQG Optimal Control [A]. *Proc IEE Conf on Control'91* [C]. Edinburgh: IEE, 1991: 768-772

- [7] Goodwin G C, Sun K S. *Adaptive Filtering Prediction and Control* [M]. Englewood Cliffs: Prentice Hall, 1984
- [8] Astrom K J, Wittenmark B. *Computer Controlled Systems* [M]. 3rd ed. Englewood Cliffs: Prentice Hall, 1997.