Optimization of classification theory and application in stochastic functional analysis

yuanxiuli

Zaozhuang kejizhiyexueyuan, shandong, tengzhou, 277500

Keywords: stochastic functional; support vector machine; convergence; data classification

Abstract. A class of stochastic functional differential equation boundary value convergence problem is researched, and the support vector machine (SVM) optimization classification theory is taken for the data classification, and the training sample set processing algorithm is proposed based on the stochastic functional analysis, the stability and convergence of the stochastic functional characteristics is improved, new algorithm can effectively absorb or eliminate redundant information and avoid additional data, repeatedly re calculation can be avoided, it can radically reduce the number of iterations and calculation cost. Stability properties of positive semidefinite minimum positive characteristic zonal sparse conditions is proved by mathematical deduction, the line approximation error continuous edge stability analysis is taken for stochastic functional value, singular decomposition is taken with conjugate gradient method, the boundary conditions are incorporated into the stochastic functional elliptic function, the small positive semidefinite minimum feature of the stiffness matrix is given, the asymptotic convergence condition of Stochastic Functional Differential Equations is obtained, and the boundary value is solved. KDD_CUP2012 database is used in simulation for data classification, simulation results show that, the classification model can classify all kinds of data and reorganize the data effectively. It can improve the data processing ability, guarantee the convergence of stochastic functional differential equations boundary value.

Introduction

In recent years, with the development of network technology, a large amount of data need for transmission constantly, we need to establish a mathematical model and algorithm, and realize the optimization of the classification of the data effectively, the problem that how to cluster the massive data and massive information for the data analysis is the key step and key base of the data mining and information processing, classification data is mainly faced to the original data, a certain algorithm is designed for data classification and clustering, it is a topic worthy of in-depth study. By clustering or classification processing, it is the important procedure in the fault diagnosis subject and target recognition subject based on the clustering and classification, the similar characteristics of the data will be gathered together, and exit from an output of the classifier[1-3].

The traditional methods of data classification contain neural network classifier, fuzzy C mean classifier, SVM classifier, and so on, in which the SVM classifier is most typical, these methods are effective data classification methods, and they are widely applied in pattern recognition, image processing and data mining fields. The convergence and iteration number and the amount of calculation problem of SVM is the key to solve the mathematical classification problem. of SVM optimization classification theory is analyzed[4], and adding reduction type support vector machine training sample set processing algorithm is proposed in this paper for improving the stability and convergence of the solution of the stochastic functional characteristics. The model of the stochastic functional differential problem was first proposed by Vapnik et al. Stochastic functional differential equations boundary value is researched, disturbance characteristic functional convex combination model becomes effective tool of analysis the continuity, stability and resolution. The stability and convergence of stochastic functional differential equation model is analyzed, to solve the data classification, and the problems such as linear regression model and statistical signal processing structure and stable solution can be solved[5].

Therefore, in this paper, the support vector machine (SVM) optimization classification theory is taken for the data classification, and the training sample set processing algorithm is proposed based on the stochastic functional analysis, stability properties of positive semidefinite minimum positive characteristic zonal sparse conditions is proved by mathematical deduction, the line approximation error continuous edge stability analysis is taken for stochastic functional value, singular decomposition is taken with conjugate gradient method, the boundary conditions are incorporated into the stochastic functional elliptic function, the small positive semidefinite minimum feature of the stiffness matrix is given, the asymptotic convergence condition of Stochastic Functional Differential Equations is obtained, and the boundary value is solved[6]. KDD_CUP2012 database is used in simulation for data classification, simulation results show that, the classification model can classify all kinds of data and reorganize the data effectively. It can improve the data processing ability, guarantee the convergence of stochastic functional differential equations boundary value. It will have good application value in data analysis.

Basic knowledge and stochastic functional theory

Stochastic functional differential equation and foundation of boundary value problem. The stability problem of stochastic functional differential equations boundary value is reached, and the stochastic functional differential equation need to be constructed firstly, the mathematical model method is used for solving nonlinear ordinary differential equations, nonlinear partial differential equations, the control algorithm is designed, firstly gives the stochastic functional nuclear in complex differential equation model[7], in the stochastic functional core, a stable system mathematical model of stochastic functional matrix is constructed, Hausdorff expression $u(\xi)$ is expanded as a stochastic functional elliptic sine function $sn\xi$, the series of mathematical model based on Hausdorff matrix is defined as:

$$u = \sum_{j=0}^{n} a_j sn^j \xi \tag{1}$$

In Cauchy kernel, the highest order mathematical model of stochastic functional matrix in the generalized matrix distribution is defined as:

$$O(u(\xi)) = n \tag{2}$$

According to the equilibrium theory in translation and expansion, considering the unconstrained optimization of stochastic functional differential equation model:

$$u_{tx} + qu_{yy} + 6(uu_x)_x + u_{xxx} = 0 \tag{3}$$

While $q = 2$, Banach space of differential equation is represented as (2+1) - dimensional differential equations, the constraint condition is:

$$\frac{du}{d\xi} = \sum_{j=0}^{n} ja_j sn^{j-1}\xi cn\xi dn\xi \tag{4}$$

The singular decomposition is taken by conjugate gradient method, multi-objective optimization local search method is used to solve the solution of the above equations, the results are taken into the stochastic functional elliptic function, while $m \to 1$, $sn\xi \to \tanh\xi$, generalized central symmetric matrix is constructed for sparse symmetric positive semi definite band of stiffness degradation for the initial estimation:

$$u = \sum_{j=0}^{n} a_j \tanh^j \xi \tag{5}$$

An autoregressive linear optimal solution set is obtained, according to the theory of multi-objective optimization, stochastic functional complex variable function is determined by n decision variables and , m target variables, said its mathematical model is expressed as:

$$\begin{aligned} \min \quad & F(x) = (f_1(x), f_2(x), ..., f_m(x))^T \\ s.t. \quad & g_i \leq 0, \qquad i = 1, 2, ..., q \\ & h_j = 0, \qquad j = 1, 2, ..., p \end{aligned} \tag{6}$$

Lemma 1: In finite domain $Y^H(q)A(0)$, the n order matrix A is the convex combination model, sufficient condition of convex matrix A is the matrix of arbitrary order sub matrix is non singular.

The disturbance characteristic functional is a kind of a neighborhood features control model for the origin is:

$$\begin{cases} \dot{m}_i(t) = -a_i m_i(t) + b_i(p_i(t-\sigma), p_2(t-\sigma), \cdots, p_n(t-\sigma)) \\ \dot{p}_i(t) = -c_i p_i(t) + d_i m_i(t-\tau) \end{cases} \quad (7)$$

In the formula, turbulence characteristics of average type for stochastic functional analytic matrix equation is reversible convex combination of analytical matrix, therefore, it needs to give the turbulence characteristics of functional principle, according to the autocorrelation function information vector iteration model, when the autocorrelation vector the minimum mutual information is achieved asymptotic stability, then global asymptotic stability is obtained.

Boundary value solution and stability characteristics. The convex combination of subspace neighbourhood $N(0)$ exists, the mutual information of average disturbance characteristic functional $H(\bullet)$ is set, the disturbance characteristics connection weights for the system function is:

$$H(s) + Y(s) = G_m(s)U(s) \quad (8)$$

When the controlled object and a state feedback controller is limited with learning constraints, the system reaches equilibrium, the characteristic functional function control model is $\dot{x}(t) = Ax(t) + Bu(t)$ and $u_c(t) = Kx_c(t)$, the closed-loop system is obtained, which is:

$$\dot{x}(t) = Ax(t) + BKx(t - d_s(t) - d_a(t)) \quad (9)$$

This paper uses the stochastic functional to solve the stable solution of differential equation, the learning algorithm of the controlled object is used for computing the real-time adaptive eigenvalue decomposition, expressed as:

$$f(\mathbf{x}, q) = \sum_{i=1}^{m} \{\frac{1}{q} \ln\{\exp[qf_i(\mathbf{x})] + \exp[-qf_i(\mathbf{x})]\}\}^P \quad (10)$$

Assumed arbitrary initial matrix of $f(\mathbf{x}), f(\mathbf{x}, q)$ in range $\mathbf{x}$ is $f(\mathbf{X}), f(\mathbf{X}, q)$, because $F(\mathbf{X})$ is the only minimum norm solution of $f(\mathbf{x})$, therefore:

$$\overline{f}_q^{(0)} = \overline{f}_q = f(m(\mathbf{X}^{(0)}), q^{(0)}) \quad (11)$$

On the basis of stochastic functional convergence conditions, the functional disturbance characteristics are extracted, and the construction disturbance characteristic functional linear equations are constructed as:

$$\sup F(\mathbf{X}, q) - \sup f(\mathbf{X}, q)$$
$$= \sum_{i=1}^{m} \{\frac{1}{q} \{\exp[q \sup F_i(\mathbf{X})] + \exp[q \sup(-F_i(\mathbf{X}))]\}\}^P$$
$$- \sum_{i=1}^{m} \{\frac{1}{q} \ln\{\exp[q \sup f_i(\mathbf{x})] + \exp[q \sup(-f_i(\mathbf{x}))]\}\}^P$$
$$= \sum_{i=1}^{m} \frac{p}{q^p} \xi_i^{p-1} \{\exp[q \sup F_i(\mathbf{X})] + \exp[q \sup(-F_i(\mathbf{X}))]\}$$
$$- \ln\{\exp[q \sup f_i(\mathbf{x})] + \exp[q \sup(-f_i(\mathbf{x}))]\}\}$$
$$= \frac{p}{q^p} \xi_i^{p-1} \sum_{i=1}^{m} \frac{\exp[q\lambda_1][\sup F_i(\mathbf{X}) - \sup f_i(\mathbf{x})]}{\exp[q\lambda_1] + \exp[q\lambda_2]} +$$
$$\frac{\exp[q\lambda_2][\sup(-F_i(\mathbf{X})) - \sup(-f_i(\mathbf{x}))]}{\exp[q\lambda_1] + \exp[q\lambda_2]}$$
$$\leq \frac{p}{q^p} \xi_i^{p-1} \sum_{i=1}^{m} \frac{\exp[q\lambda_1]K_1 \|W(\mathbf{X})\|_\infty^2}{\exp[q\lambda_1] + \exp[q\lambda_2]} + \frac{\exp[q\lambda_2]K_1 \|W(\mathbf{X})\|_\infty^2}{\exp[q\lambda_1] + \exp[q\lambda_2]}$$
$$\leq \frac{p}{q^p} \xi_i^{p-1} m K_1 \|W(\mathbf{X})\|_\infty^2 = \frac{K}{2} \|W(\mathbf{X})\|_\infty^2 \quad (12)$$

In the interval variable with the same functional, similarly we can get:

$$\inf F(\mathbf{X}, q) - \inf f(\mathbf{X}, q) \leq \frac{K}{2} \|W(\mathbf{X})\|_\infty^2 \quad (13)$$

Because the function $\ln x, e^x$ is monotone increasing function, so:

$$W(F(\mathbf{X}, q)) - W(f(\mathbf{X}, q)) \leq K \|W(\mathbf{X})\|_\infty^2 \quad (14)$$

Given the stiffness matrix of small positive semidefinite minimum feature as:

$$\Delta E = -\eta \left[\left(\frac{\partial E}{\partial \omega} \right)^2 + \left(\frac{\partial E}{\partial b} \right)^2 \right] \tag{15}$$

If $C_o(x^*) = 0$, then:

$$Y(P,Q,\beta) = Y\left[red(P,Q,\beta),Q,\beta \right] \tag{16}$$

The convergence problem of convex combination model is equivalent to range expansion problem of $F(\mathbf{x},q)$, and it has at least two order of convergence, so to construct a convex combination model of Schur convergence disturbance characteristics based on functional theory, and it is convergent.

Optimal classification model based on stochastic functional analysis

Based on the above analysis, boundary value convergence problem of this kind of stochastic functional differential equation is analyzed, optimization classification theory of SVM is taken, by adding reduction type support vector machine training sample set processing algorithm, to improve the stability and convergence of the solution of the stochastic functional characteristics. The standard solution of support vector machine is expressed as:

$$\min_{0 \le \alpha_i \le c} W = \frac{1}{2} \sum_{i,j=1}^{l} y_i y_j \alpha_i \alpha_j K(x_i,x_j) - \sum_{i=1}^{l} \alpha_i + b \left(\sum_{i=1}^{l} y_j \alpha \right) \tag{17}$$

Where, (x_i,x_j) is the data sample, b is the Lagrange operator, sample set is:

$$S = \left\{ (x_1,x_1), \cdots, (x_l,x_l) \right\} \tag{18}$$

According to the convex optimization KKT theorem, the formula (18), the first-order derivative is solved, the standard support classification KKT condition for convex optimization is available:

$$G_i = \sum_j \alpha_j y_i y_j K(x_i,x_j) + y_i b - 1 \tag{19}$$

KKT convex optimization condition is expressed as:

$$G_i = \begin{cases} \ge 0, \alpha_i = 0 & S_R \\ = 0, 0 < \alpha_i < C & S_S \\ \le 0, \alpha_i = C & S_E \end{cases} \tag{20}$$

Where, $\sum_{i=1}^{l} y_i \alpha_i = 0$, S_R, S_S and S_E are three sub sets of SVM training set, respectively they are: S_R — the redundant vector set; S_S -- margin support vector set; S_E -- -- wrong classification support vector set. In order to effectively obtain increase adjustment training sample set of optimal features, based on learning algorithm, need to build rules to update the corresponding increase or decrease, so that after the adjustment of the sample set, it is still satisfy the convex optimization conditions of standard support vector machine, asymptotic stability is analyzed, and $\mathbf{x}^{(0)}$ is divided into $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(N)}$, $\mathbf{x}^{(0)} = \bigcup_{i=1}^{N} \mathbf{x}^{(i)}$, set:

$$W_N = \max_{1 \le i \le N} \left\| W(\mathbf{x}^{(i)}) \right\|_{\infty} \tag{21}$$

Assumed $\{q_N\}$ is monotonically increasing, $q_N \ge 1$ and when $N \to \infty$, get $q_N \to \infty$, after the limited step length calculation, the disturbance feature functional is:

$$\alpha_N = \min_{1 \le i \le N} \{ \inf F(\mathbf{x}^{(i)}, q_N) \} \tag{22}$$

the line approximation error continuous edge stability analysis is taken for stochastic functional value, the control function is:

$$\dot{x}(t) = Ax(t) + Bx(t - d_1(t) - d_2(t))\sqrt{2}, x(t) = \phi(t) \quad t \in [-h,0] \tag{23}$$

In order to realize the optimization classification in the SVM, input new vectors into the support vector machine:

$$x(t) = (x_0(t), x_1(t), \cdots, x_{k-1}(t))^T \tag{24}$$

For $x(t)$, calculate the Euclidean distance to the output nodes with weighted vector, it can be expressed as:

$$d_j = \sum_{i=0}^{k-1} (x_i(t) - \omega_{ij}(t))^2 \ , \ j = 0,1,\cdots,N-1 \tag{25}$$

Where $\omega_j = (\omega_{0j}, \omega_{1j}, \cdots, \omega_{k-1,j})^T$, $\forall \varepsilon > 0, \exists \hat{N} > 0$, $\overset{\underline{\underline{\,}}}{=}\ N > \hat{N}$, $\left| \min\limits_{x \in X^{(0)}} e = (f(\mathbf{x}) - \alpha_N) \right| < \varepsilon$, the boundary conditions are incorporated into the stochastic functional elliptic function, according to the set S_s, sample add 1, the set mobiles with marginalization. The optimization output is:

$$Q^-.R^- = Q^-.R_{tt} - \frac{Q^-}{R_{tt}}(R_{*t}.R_{*t})_{\backslash tt} = Q^-.R_{tt} + (R_{*t}.R_{*t})_{\backslash tt} = I \tag{26}$$

In the phase space, when the dimension is added from m to $m+1$, the distance is shown as:

$$R_{(m+1)n} = \left\| X_{\eta(n)} - X_n \right\|_2^{(m+1)} = \sqrt{\sum_{l=0}^{m} \left(x_{\eta(n)+l\tau} - x_{n+l\tau} \right)^2} = \sqrt{\left[\left\| X_{\eta(n)} - X_n \right\|_2^{(m)} \right]^2 + \left[x_{\eta(n)+m\tau} - x_{n+m\tau} \right]^2} \tag{27}$$

Where, $\langle Q_s \rangle$ is shown as the mean value of judge statistical quantity. According to the s value which determines classification and nonlinearity of the data series, $\alpha = 5\%$, if the difference value of Q_0 and $\langle Q_s \rangle$ overcome a critical value Q_c, make $p(|Q_0 - \langle Q_s \rangle| > Q_c) \leq 0.05$, the standardized normal distribution meets the condition:

$$\begin{aligned}
0 &\leq \left| \min_{x \in X^{(0)}} (f(\mathbf{x}, q_k) - \alpha_N) \right| \\
&= \left| \min_{x \in X^{(0)}} (f(\mathbf{x}, q_k) - \inf F(\mathbf{X}^{(k)}, q_k) \right| \\
&= \left| \min_{x \in X^{(k)}} (f(\mathbf{x}, q_k) - \inf F(\mathbf{X}^{(k)}, q_k) \right| \\
&\leq K \left\| W(\mathbf{X}^{(k)}) \right\|_\infty
\end{aligned} \tag{28}$$

While, $\left\| W(\mathbf{X}^{(k)}) \right\|_\infty \to 0$, sandwich theorem by stochastic functional differential equation with perturbation is：

$$\left| \min_{x \in X^{(0)}} (f(\mathbf{x}, q_k) - \alpha_N) \right| \to 0 \tag{29}$$

Then, $\forall \varepsilon > 0, \exists N_1 > 0$, while $N > N_1$, in the subspace, the boundary value solution:

$$\left| \min_{x \in X^{(0)}} (f(\mathbf{x}, q_k) - \alpha_N) \right| < \frac{\varepsilon}{2} \tag{30}$$

While $f(\mathbf{x}, q_k) \to f(\mathbf{x})$, that $\exists N_2 > 0$, then:

$$\left| \min_{x \in X^{(0)}} (f(\mathbf{x}, q_k) - f(\mathbf{x})) \right| < \frac{\varepsilon}{2} \tag{31}$$

Therefore, take $\hat{N} = \max\{N_1, N_2\}$, while $N > \hat{N}$, then:

$$\begin{aligned}
&\left| \min_{x \in X^{(0)}} e = (f(\mathbf{x}) - \alpha_N) \right| \\
&= \left| \min_{x \in X^{(0)}} [(f(\mathbf{x}, q_k) - \alpha_N) - (f(\mathbf{x}, q_k) - f(\mathbf{x}))] \right| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned} \tag{32}$$

Thus, the singular decomposition is taken with conjugate gradient method, the boundary conditions are incorporated into the stochastic functional elliptic function, and stochastic functional differential equations boundary value of convex combination model is convergent.

Numerical simulation test

Based on the MATLAB software, the simulation is taken for test the performance of the algorithm in data classification. Stochastic functional optimization classification is taken for the kernel function as above, and the KDD_CUP2012 database is used in simulation for data classification, 3 data sets are taken as the samples, collecting 3 data sets, respectively, the amount of data are 50, 100, 150, and 200200, 200 times number of iterations are taken in simulation, validate the algorithm

in the realization of data classification, mathematical model for data classification is constructed, data set classification results are shown in Figure 1.

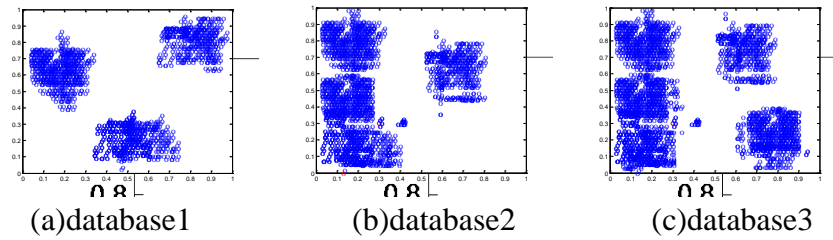(a)database1   (b)database2   (c)database3

Figure 1 Data set classification results

From figure 1, it shows that the method has good classification performance. In order to compare the performance, the traditional FCM is taken in simulation, and the accuracy of classification is taken as the evaluation index, the results are shown in Table 1. It show that the accuracy of classification improved by 20.8%, it has better performance.

Table 1 Accuracy of classification

Data sets	FCM	SVM stochastic functional classification algorithm
Database1	0.8224	0.94231
Database2	0.8442	0.91678
Database3	0.8565	0.97724

Conclusions

In this paper, a class of stochastic functional differential equation boundary value convergence problem is researched, and the support vector machine (SVM) optimization classification theory is taken for the data classification, and the training sample set processing algorithm is proposed based on the stochastic functional analysis, the stability and convergence of the stochastic functional characteristics is improved, new algorithm can effectively absorb or eliminate redundant information and avoid additional data, the small positive semidefinite minimum feature of the stiffness matrix is given, the asymptotic convergence condition of Stochastic Functional Differential Equations is obtained, and the boundary value is solved. KDD_CUP2012 database is used in simulation for data classification, simulation results show that, the classification model can classify all kinds of data and reorganize the data effectively. It can improve the data processing ability, guarantee the convergence of stochastic functional differential equations boundary value. The research has good application value.

References

[1] DYO V, ELLWOOD S A., MACDONALD D W, et al. WILDSENSING: design and deployment of a sustainable sensor network for wildlife monitoring[J]. ACM Trans.on Sensor Networks, 2012, 8(4):1-33.
[2] SUN Huijun, ZHANG Hui, WU Jianjun. Correlated scale-free network with community: modeling and transportation dynamics [J]. Nonlinear Dynamics, 2012, 69(4): 2097-2104.
[3] Yick J, Mukherjee B, Ghosal D．Wireless sensor network surwey [J]. Computer Networks, 2008, 52(12): 2292-2330.
[4] Joachims T, Finley T, Yu CNJ. Cutting-Plane training of structural SVMs[J]. Machine Learning, 2009,77(1): 27−59.

[5] Terrill Philip lan, Wilson Stephen James, Suresh Sadasivam, et al. Attractor structure discriminates sleep states: Recurrence plot analysis applied to infant breathing patterns [J]. IEEE Transactions on Biomedical Engineering, 2010, 57(5): 1108-1116.

[6] Guhathakurta Kousik, Bhattacharya Basabi, Chowdhury A. Roy. Using recurrence plot analysis to distinguish between endogenous and exogenous stock market crashes[J]. Physica A: Statistical Mechanics and its Applications, 2010, 389(9): 1874-1882.

[7] Dunn J C. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. Q.cybernet, 1974(3): 32-571.