TELKOMNIKA, Vol.11, No.11, November 2013, pp. 6406~6412
e-ISSN: 2087-278X

Multi-Objective Optimization Algorithms Design based on Support Vector Regression Metamodeling

QiZhang*, Guangchun Yang, Rui Tang, Guoqi Xiang
Panzhihua University, 10# Airport Road, Panzhihua City, Sichuan Province
P.R.China, 617000, +86-812-3371000
*Corresponding author, e-mail: pzhuzq@126.com

Abstract

In order to solve the multi-objective optimization problem in the complex engineering, in this paper a NSGA-II multi-objective optimization algorithms based on support vector regression metamodeling is presented. Appropriate design parameter samples are selected by experimental design theories, and the response samples are obtained from the experiments or numerical simulations, used the SVM method to establish the metamodels of the objective performance functions and constraints, and reconstructed the original optimal problem. The reconstructed metamodels was solved by NSGA-II algorithm and took the structure optimization of the microwave power divider as an example to illustrate the proposed methodology and solve the multi-objective optimization problem. The results show that this methodology is feasible and highly effective, and thus it can be used in the optimum design of engineering fields.

***Keywords**: NSGA-II, metamodeling, optimization design, support vector machine*

1. Introduction

In the practical optimizing design, there is an implicit function relationship between the objective function and the design variables of the problem to be optimized, and its expression is often high linearity, multi-parameters and so forth. If we directly use the existing optimization method to optimize them, there will be some optimization problems. The traditional method mainly relies on the experience to call numerical simulation model for the calculation repeatedly. There are more shortcomings about them, such as time-consuming, inefficient, and difficult to get a global optimal.

Response surface, Kriging model and neural network are commonly used metamodeling in current study [1-3]. The response surface method is easy to implement, but the ability to approximate nonlinear problems is poor. The Kriging is higher accuracy for nonlinear problems, but to acquire and use the model is difficult. The Neural network exists over learning phenomenon. The model accuracy and generalization ability depends on its structure and a lot of learning samples, and the selection of the structure lacks of theoretical guiding.

The support vector machine (SVM) [4-6] is based on statistically VC dimension theory and structural risk minimization principle. The practical problems of the small sample, nonlinearity, high dimension and local minima points are solved by using it. And it has strong generalization ability. The SVM is an effective method to establish the nonparametric regression model of the small sample problem.

NSGA-II algorithm, proposed in 2002, is an improved version of NSGA [7-8], which is high efficiency, convergence. Optimal solutions which meet the different requirements can be found in a single optimization process.

In this paper, a NSGA-II multi-objective optimization algorithms based on Support Vector Regression Metamodeling is presented. Used the SVM method to establish the metamodels of the objective performance functions and other constraints, and reconstructed the original optimal problem. The reconstructed metamodels was solved by NSGA-II algorithm.

Received May 9, 2013; Revised August 3, 2013; Accepted August 16, 2013

2. Support Vector Regression

Assumed, training samples $\{(x_1, y_1), (x_2, y_2) \cdots, (x_n, y_n)\}$, $x_i, y_i \in \mathrm{R}$ are known. We use the following linear regression function to fit them.

$$f(x) = \omega \cdot x + b \tag{1}$$

In order to get a good fitting effect, assume that all the training samples can be fitted linearly under the accuracy of ε . Then, regression estimation function is equivalent to look for a minimum of $\|\omega\|$, it can be expressed as convex optimization problem.

$$\begin{cases} \text{Min} & \frac{1}{2}\|\omega\|^2 \\ \text{S.T.} & \begin{cases} \omega \cdot x_i + b - y_i \le \varepsilon \\ y_i - \omega \cdot x_i - b \le \varepsilon \end{cases} \end{cases} \tag{2}$$

Take allowed fitting error into account, positive slack variables ξ_i, ξ_i^* is taken in. So the fitting problem is transformed into the following optimization problem.

$$\text{Min} \quad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{3}$$

$$\begin{aligned} & y_i - \omega \cdot x_i - b \le \varepsilon + \xi_i \\ \text{S.T.}\ & \omega \cdot x_i + b - y_i \le \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \ge 0 \quad i = 1,2,\cdots n \end{aligned} \tag{4}$$

C is a regularized constant determining the trade-off between the training error and the model flatness. ε can be viewed as a tube size equivalent to the approximation accuracy in training data. By using Lagrange multipliers techniques, this optimization formulation can be transformed into the dual problem, and its maximal dual function in Equation (3) and Equation (4) which has the following form:

$$\begin{aligned} &\text{Max} \left\{ W(\alpha,\alpha^*) = -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\langle x_i, x_j\rangle + \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)y_i - \sum_{i=1}^{n}(\alpha_i + \alpha_i^*)\varepsilon \right. \\ &\qquad \text{S.T.} \begin{cases} \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_i, \alpha_i^* \le C \quad i = 1,2,\cdots n \end{cases} \end{aligned} \tag{5}$$

Where α_i, α_i^* are so-called Lagrange multipliers.

SVR linear fit function, given by Equation 3 and Equation 4, has the following explicit form.

$$\begin{aligned} f(x) &= \omega \cdot x + b \\ &= \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\langle x, x_i\rangle + b \end{aligned} \tag{6}$$

If the training samples can't be fitted linearly, we can adopt the following ideas to solve it. Firstly, the training samples are mapped into high dimensional feature space by using nonlinear transformation, Secondly, we use the previous method to fit them linearly in the high-dimensional feature space, the fitting effect is the same as being fitted in the original space. By the functional theory, Inner-Product in the high dimensional feature space is substituted by a kernel function

($K(x, x^{'})$) in the original space. So the Support vector regression nonlinear fitfunction is obtained:

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x, x^{'}) + b \quad (7)$$

2.1. SVR Parameters Selection

The value of the SVR parameters (C, σ and ε) are relevant to the learning samples and practical issues. Parameter C determines the trade-off cost between minimizing the model's complexity and minimizing the training error. Parameter ε controls the width of the ε -insensitive zone, used to fit the training data. It is equivalent to the approximation accuracy placed on the training data points. The value of ε can affect the number of SVs used to construct the regression function. Larger ε value result in fewer SVs selected, and result in more 'flat' (less complex) regression estimates. Bandwidth of the kernel function σ: Kernel parameter σ is selected to reflect the input range of the training/test data. For multivariate problems, the range of parameters C, ε and σ investigated in [9-10] are set to the following form:

$$\varepsilon \in [0, 5\bar{\sigma}], \bar{\sigma}^2 = \frac{kl^{1/5}}{kl^{1/5} - 1} \sum_{i=1}^{l} (y_i - \hat{y}_i)^2 \quad (8)$$

$$C_{\min} = \max(|\bar{y} + 3\sigma|, |\bar{y} - 3\sigma|) \quad (9)$$

Where, k=3, l denotes the sample size, $\bar{y}$ and $\bar{\sigma}$ are the mean and standard deviation of y values of training data. $\hat{y}_i$ is the output of SVR, when ε =0.

3. Non-dominated Sorting Genetic Algorithm-NSGAII

NSGA-II algorithm is high efficiency, convergence. Optimal solutions which meet the different requirements can be found in a single optimization process, Figure 1 shows the algorithm flow of NSGA-II.

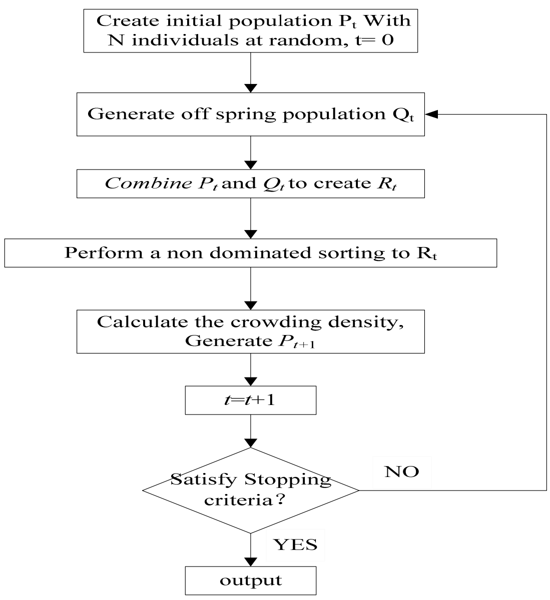

Figure 1. The Algorithm Flow of NSGA-II

Then NSGA-II algorithm can be described as following.

Step1: Create initial population P_t, with N individuals at random, which P_t is current population.

Step2: Randomly choose two individuals as parents and use crossover and mutation operator to get two new individuals. Repeat above selection, crossover and mutation process till N new individuals are created. These N new individuals compose population Q_t.

Step3: Combine the population Pt and Qt to get bigger population Rt.

Step4: Perform a non-dominated sorting to Rt.

Step5: Calculate the crowding density of all individual with rank numbers. Put right individuals into population Pt+1 according to descending order of density estimation of individuals till the number of individuals in population Pt+ 1 is exactly N.

Step6: Set P_{t+1} as current population P_t and go to step two if stop condition doesn't been reached.

4. SVR-NSGAII Algorithm

Generally, a multi-objective model has following form:

$$\min F(X) = [f_1(X), f_2(X), \cdots, f_i(X), \cdots, f_n(X)]$$
$$\text{s.t.}\begin{cases} g_j(X) \le 0, j = 1,2,\cdots,l \\ h_j(X) = 0, j = 1,2,\cdots,m \\ x_{j\min} \le x_j \le x_{j\max}, j = 1,2,\cdots,k \end{cases} \tag{10}$$

Where, $f_i(x)$ denotes the ith objective function, $g_j(x)$ denotes the ith constraint, x_{jmin} and x_{jmax} are the jth dimension design variable bound. Generally, objective functions and constraints are implicit mathematics forms, and are usually obtained by simulating or experiment. In order to improve the optimization efficiency multi-objective model are replaced by metamodels, multi-objective model is further formulated as following form:

$$\min F = [Y_1^*, Y_2^*, \cdots, Y_i^*, \cdots, Y_n^*], Y_i^* = f_i(x) \quad i = 1,2,\cdots n$$
$$\text{s.t.} \begin{matrix} \{G_j^* = g_j(x)\} \le 0 & j = 1,2,\cdots,l \\ \{H_j^* = h_j(x)\} = 0 & j = 1,2,\cdots,m \\ x_{j\min} \le x_j \le x_{j\max} & j = 1,2,\cdots k \end{matrix} \tag{11}$$

Where, Y_i^* , G_j^* and H_j^* denote the metamodels of the i^th objective function and the j^th constraint.

In this paper, the SVM method was used to establish the metamodels of the ith objective functions and the jth constraints. The reconstructed metamodels was solved by NSGA-II algorithm. Figure 2 shows NSGAII optimization algorithm flow based on SVR.

Step 1: Experiment Design

Select variables with a significant impact as a design variable, choice test design method, design training samples.

Step 2: Optimize SVR parameters

Step 3: Construction SVR metamodels

Approximation of the objective function and constraint functions, construct support vector regression metamodels by training samples and optimized parameters. If the constructed metamodels' approximation accuracy does not meet the requirements, takes the optimization results of each iteration as a new training sample, update model, improve its accuracy.

Step 4: Multi-objective optimization

Take the NSGA-II algorithm to optimize the reconstructed metamodels,

Step 5: According to actual requirements, select the satisfactory solutions. Use distribution uniformity and diversity of Pareto set to evaluate the quality of Pareto set. If the Pareto set can satisfy the requirement, output the set, otherwise, back to STEP3.

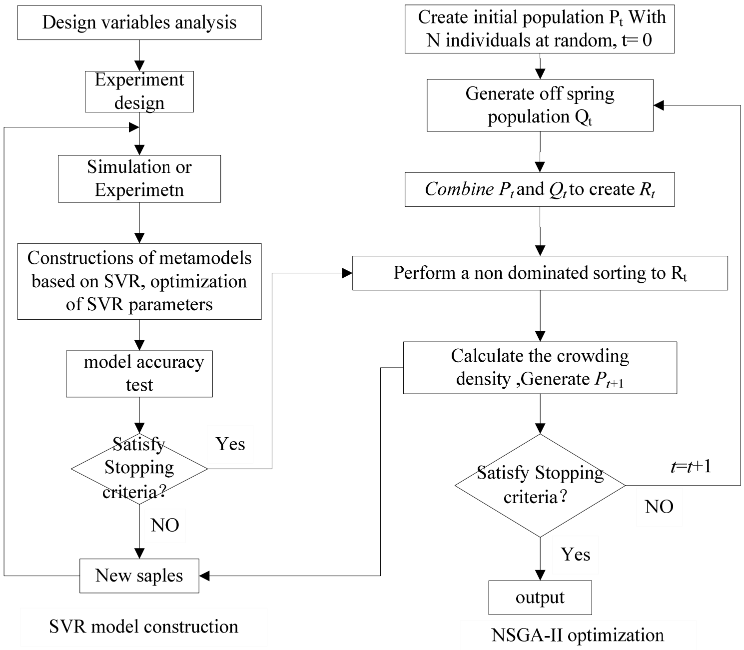

Figure 2. SVR-NSGAII

5. Experiment

For demonstrating the feasibility of the method mentioned above, the structure optimization of the microwave power divider is adopted as an example to illustrate this method, Figure 3 shows the structure of power divider. In this example, the goal of structure optimization is that magnitude ratio, phase difference and VSWR between two ports of the microwave power divider equal to 0.5, 0 and 1 respectively. Design variables size is ten, and their value range are showed in Table 1.

Table 1. Value Domain of Design Variables

Variable name	Min/mm	Max/mm
L	6.8	9.2
R	6.8	9.2
H2	2.6	2.7
H1	0.48	0.72
R1	0.15	0.25
V1	0.75	0.85
R2	0.38	0.62
V2	0.48	0.72
A	5.5	6.5
B	1	2

Take the NSGA-II to optimize the examplementioned, choose the floating-point encoding as the encoding way, the algorithm population size is 30, and the evolution algebra is 50, crossover probability is 0.9, mutation probability is 0.01, Pareto set is shown in Table 2.

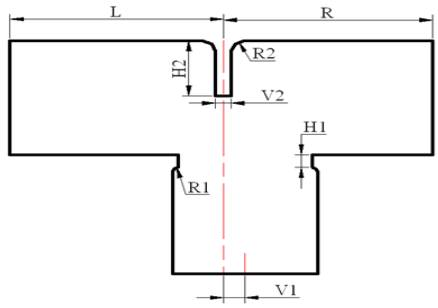

Figure 3. The Structure of the Power Divider

Table 2. Pareto Set by using SVR-NSGAII

Order	VSWR (Fitness1)	magnitude ratio (Fitness2)	phase difference (Fitness3)
1	0.34041	0.24929	-2.454
2	0.27776	0.18807	4.119
3	0.32864	0.22451	6.463
4	0.3155	0.21844	-5.442
5	0.040855	0.008407	12.2048
6	0.14775	0.095993	5.774
7	0.11352	0.027649	-11.946
8	0.27724	0.21075	-5.171
9	0.16466	0.12502	10.117
10	0.2189	0.17558	-8.524
11	0.29698	0.18111	7.421
12	0.27832	0.21418	-3.268
13	0.35032	0.26497	1.017
14	0.24809	0.18931	6.8732
15	0.32258	0.24472	-2.066
16	0.063233	0.028701	-11.8112
17	0.22864	0.14101	-5.092
18	0.23938	0.17153	4.845
19	0.18874	0.10838	8.6018
20	0.18504	0.1089	-8.201
21	0.17581	0.09145	3.953
22	0.23763	0.1694	-5.2
23	0.18996	0.14147	6.8637
24	0.25995	0.18218	4.663
25	0.16841	0.09998	-3.714
26	0.12161	0.057939	9.5752
27	0.27542	0.20982	2.464
28	0.2024	0.12938	-4.407
29	0.2865	0.19012	3.974
30	0.088151	0.035294	10.983

According to actual requirements, we can select the satisfactory solutions From Table 2.

6. Conclusion

The paper presents a NSGA-II genetic optimization design methodology based on the SVR and applied it to optimize the structure of the microwave power divider. From the results, the following conclusion can be obtained:

1. Appropriate design parameter samples are selected by experimental design theories, and the response samples are obtained from the experiments or numerical simulations, used the SVM method to establish the metamodels of the objective performance functions and constraints, and reconstructed the original optimal problem. The reconstructed metamodels was solved by NSGA-II algorithm. The proposed example results illustrate that the proposed algorithm is Accurate, efficient, and feasible

2. Designing experiment, constructing SVR response surface and NSGA-II optimization can be realized in parallel mode, this approach can significantly raise the efficiency of optimization process.

References

[1] Myers RH, Montgomery DC, Vining GG, et al. Response surface methodology: A retrospective and literature survey. *Journal of Quality Technology.* 2004: 36(1): 53-78.
[2] Wackerrnagel H. Multivariate geo-statistics. Heidelberg: Springer-Verlag. 1995.
[3] LiShuo, Xu Yuanming, Zhang Jun. Neural network response surface optimization design for composite stiffened structures. *Chinese Journal of Mechanical Engineering.* 2006; 42(11): 115-119.
[4] VAPN IKVN. Statistical learning theory. New York: Springer. 1998.
[5] Scholkopf B, Smola A. New support vector algorithm. *NeuralComputation*. 2000; 12(5): 1207-1245.
[6] Marcelino L, Ignac IOS. Support vector regress ion for the simultaneous learning of a multivariate function and its derivatives. *Neurocomputing*. 2005; 69(8): 42-61.
[7] Deb K, Agrawal S, Pratab A, et al. *A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA2II.* In the Proceedings of the Parallel Problem Solving from Nature VI Conference. Paris. France. 2000: 849-858.
[8] Deb K, Pratap A, Agarwal S, et al. A Fast and Elitist Multi-objective Genetic Algorit hm: NSGA2II. *IEEE Transactions on Evolutionary Computation*. 2002; 6(2): 182-197.
[9] Vladimir C, Ma YQ. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks*. 2004; 17: 113-126.
[10] Du JY, Hou YB. Parameters selection of support vector regression by genetic algorithms. *Systems Engineering And Electronics*. 2006; 28(9): 1430-1433.