11th World Congress on Structural and Multidisciplinary Optimisation
07th -12th, June 2015, Sydney Australia

Zooming in Surrogate Optimization Using Convolute RBF

Masao Arakawa[1]

[1] Kagawa University, Takamatsu, JAPAN, arakawa@eng.kagawa-u.ac.jp

1. Abstract
There are a lot of efficient methods in surrogate optimization. Convolute RBF is the one of them, and we have examined its effectiveness for years. However, in most case, if the number of design variables raise and/or the searching range extended, they have some problems in approximating functions both in local accuracy and global trends. They have the same reason of sparseness of given data with respect to its searching range, and which happens by the nature of large scale optimization. In this study, zooming technique in approximation is proposed. When we have a certain amount of data, we can divide given datum into some parts and restrict its searching range to some small area. Then in each part, we can achieve more accuracy for local approximation. Moreover, if we can restrict searching range smaller, we have more possibilities to achieve global solution within given searching range. Besides, we can have local optima for each part, which can be candidates for true global solution in the future. In this paper, we examined effectiveness of the method through numerical example.
2. Keywords: Surrogate Optimization, Convolute RBF, Data Distribution

3. Introduction

There are a lot of excellent studies on Surrogate Optimization. Kashiwamura, Shiratori and Yu had published response surface method by experimental design [1], in 1998. Todoroki and Ishikawa used D-Optimality in data distributions [2]. And Myers, Montgomery and Aderson-Cook's book[3] was one of the most contributions in this area. They are all based on quadratic form. Well distribution of initial data for its approximation form. As for Meta-Modeling, Spline[4], Kriking[5], RBF network[6] and SVR[7] have been widely used to have a better approximations rather than quadratic form. Among all these studies, EGO[8] is one of the best contribution that it add strategy for global optimization especially in sampling of new data points.
I have proposed parameter optimization in setting RBF[9], recommendation function for new data points[10], convolution of RBF[11,12], data distribution method, basis distribution method [13] and so on to have better efficiency in the past.
In surrogate optimization, we already have amount of datum. Which means we can scope approximation around some data points; most likely the best solution data. When we can zoom the searching range, we may have the following advantages: 1) as the searching range shrinks, it makes easier to find make better approximation, 2) it makes easier to optimize surrogate function, and so on.
In this paper, we are going to propose the method and examined its effectiveness through numerical example.

4. Convolute RBF

4.1. RBF networks

When we think about approximation of function $y(\mathbf{x})$ by linear summation of basis functions $h_i(\mathbf{x})$, approximate function $f(\mathbf{x})$ can be expressed as follow.

$$f(\mathbf{x}) = \sum_{i=1}^{m} w_i h_i(\mathbf{x}) \tag{1}$$

In Eq.(1) we assume that we have m basis functions. When we have p teaching data, we can calculate energy of RBF as follow

$$E = \sum_{j=1}^{p} \left(y(\mathbf{x}_j) - f(\mathbf{x}_j) \right)^2 + \sum_{i=1}^{m} \lambda_i w_i^2 \tag{2}$$

Learning for RBF means trying to minimize energy in Eq.(2) with respect to weights of w_j. Thus, we can have final solutions as follows.

$$\mathbf{w} = A^{-1} H^T \mathbf{y} \tag{3}$$

Where

$$A = (H^T H + \Lambda)$$

$$H = \begin{bmatrix} h_1(\mathbf{x}_1) & h_2(\mathbf{x}_1) & \cdots & h_m(\mathbf{x}_1) \\ h_1(\mathbf{x}_2) & h_2(\mathbf{x}_2) & \cdots & h_m(\mathbf{x}_2) \\ \vdots & \cdots & \ddots & \vdots \\ h_1(\mathbf{x}_p) & h_2(\mathbf{x}_p) & \cdots & h_m(\mathbf{x}_p) \end{bmatrix}$$

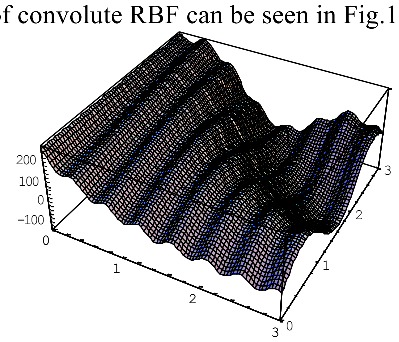
$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_m \end{bmatrix}$$

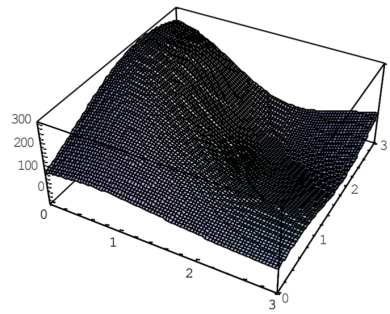
4.2. Convolute RBF

As a nature of RBF, it does not go through teaching data of $y(\mathbf{x}_i)$. However, average of error that is given from $y(\mathbf{x})$ and $y_{app}(\mathbf{x})$ is almost always close to zero. Thus, we can approximate these errors again and again until we satisfy with the error. Thus convolute RBF (Arakawa 2007)can be expressed as follow

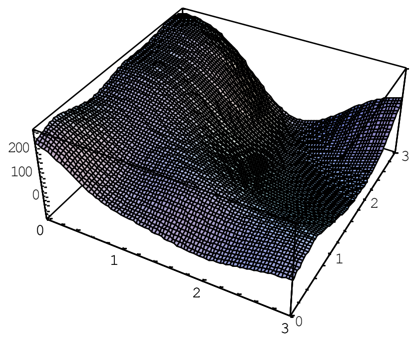
$$y_{app}(\mathbf{x}) = b(\mathbf{x}) + \sum_{l=1} f_l(\mathbf{x}) \tag{6}$$

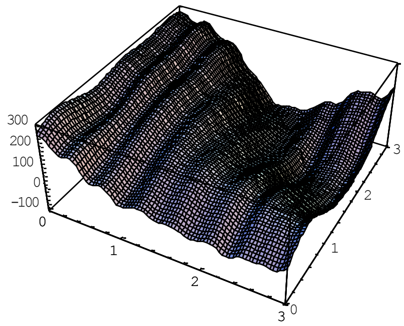
Example of convolute RBF can be seen in Fig.1.





Real function: we distribute 200 teaching data randomly   Second order of polynomial as basic function





First RBFFourth   Convolution of RBF

Fig. 1 Convolute RBF

4.3. Data Distribution

For better approximation for over all region, we would like to distribute initial data as equal as possible. For that purpose, we use 2-norm as following steps. Assume we would like to have m data.

1) Distribute m x M candidate data randomly.
2) Choose data #1 which is close to average of all m x M data.
3) Choose data #2 that has maximum distance from data #1. (Currently, we have 2 decided data)
4) Find candidate that has maximum 2-norm.

2-norm=minimum distance + second minimum distance

But, if minimum distance=0 then, 2-norm=0.

Figure 2 shows illustration of this method. At the beginning, we choose #1 as closest one to average. Then, we choose #2, that has maximum distance from #1. Then, #3 is chosen by calculating maximizing 2-norm between #1 and #2.

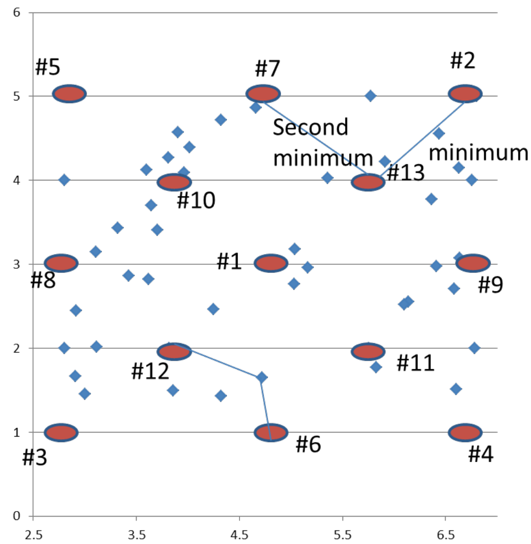


Fig. 2 Illustration of 2-norm distribution

4.4 Selection of Basis Center

For ith convolution, estimate errors at teaching data points for i-1th results. We are going to approximate these errors.

1) Find teaching data point that has maximum absolute error. ($\mathbf{x}_{max}$) Archive sign of error at this point as $sign_{max}$.
2) Find closest candidate of basis function to $\mathbf{x}_{max}$. ($\mathbf{c}_{clo}$)
3) Center of basis function is given by following

$$\mathbf{c}_j = \alpha \mathbf{x}_{\max} + (1-\alpha)\mathbf{c}_{clo} \tag{7}$$

Where, α is 0 at the first convolution and 1 at the final convolution.

4) Give radius according to the following

$$r_{ij} = ratio_k \times (\max_i - \min_i)/10 \tag{8}$$

5) Count data points within the radius. If there are more than minimum request go to 6), Otherwise enlarge radius by following
6) Count the number of data that has same sign with $sign_{max}$. If its ratio is higher than given ratio then learn RBF and estimate errors and go to 1) until it becomes number of basis function for accuracy. Otherwise go to 7)
7) Find data that has opposite sign and farther distance from center of basis function $\mathbf{c}_j$, as $\mathbf{x}_{far}$. Calculate maximum value of the following

$$t_{\max} = \max_i \left| x_{far,i} - c_{ji} \right| / (\max_i - \min_i)$$

If t_{max} is given by i_{max} variable, then change radius of i_{max} to

$$r_{i\max} = \left| x_{far,i\max} - c_{ji\max} \right|$$

Then go to 6). If we cannot satisfy both condition simultaneously for several times, we will quit there and learn RBF and go to next one.

4.6 Flow of Surrogate Optimization

Figure 3 shows the flow chart of the surrogate Optimization.

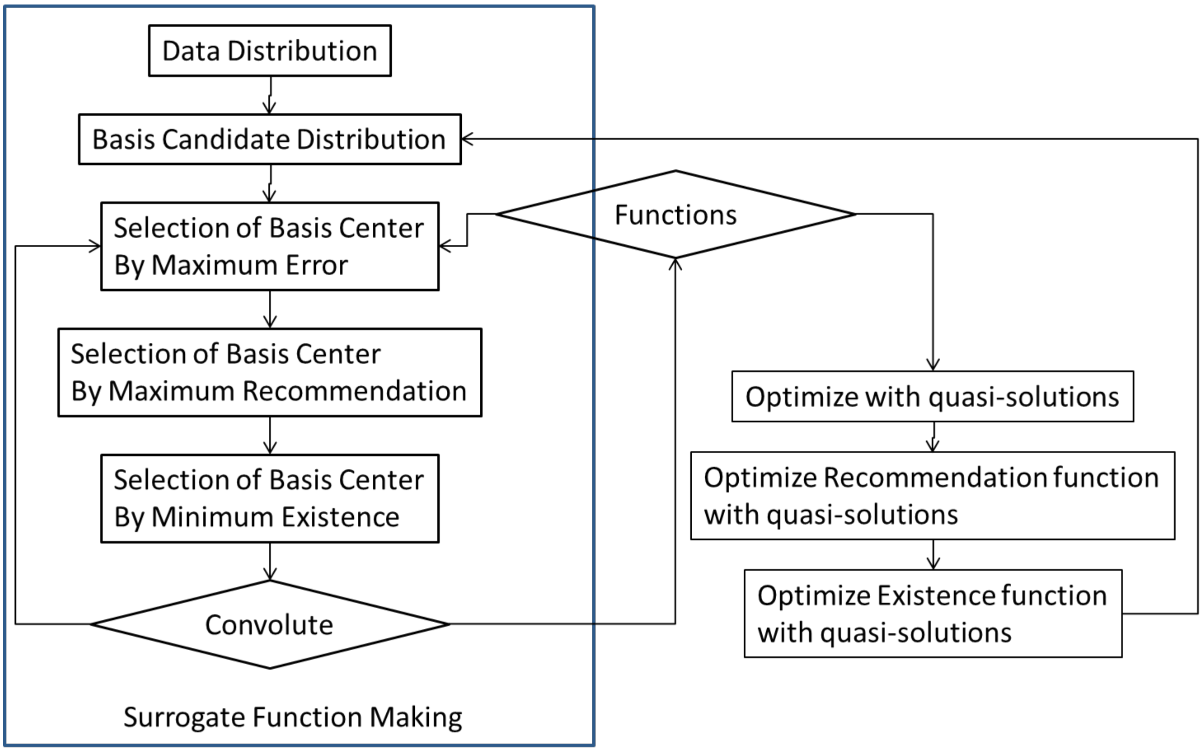


Fig. 3 Flow Chart of Surrogate Optimization

5. Prosed Method

When the number of design variables increase and/or searching range for each design variables expanded, it is sometimes difficult to obtain global optimization within a small number of function calls. Even in those cases, as a nature of surrogate optimization, we can get close to global optimization relatively in a small number of function calls. From there on, we need to repeat a number of iterations to find the solutions. One of the ways to get rid of these situations is to zoom the searching range with existence information. In this study, we propose to shrink the searching range close to the best data when it gets some amount of datum. Figure 4 shows the flow of the proposed method.

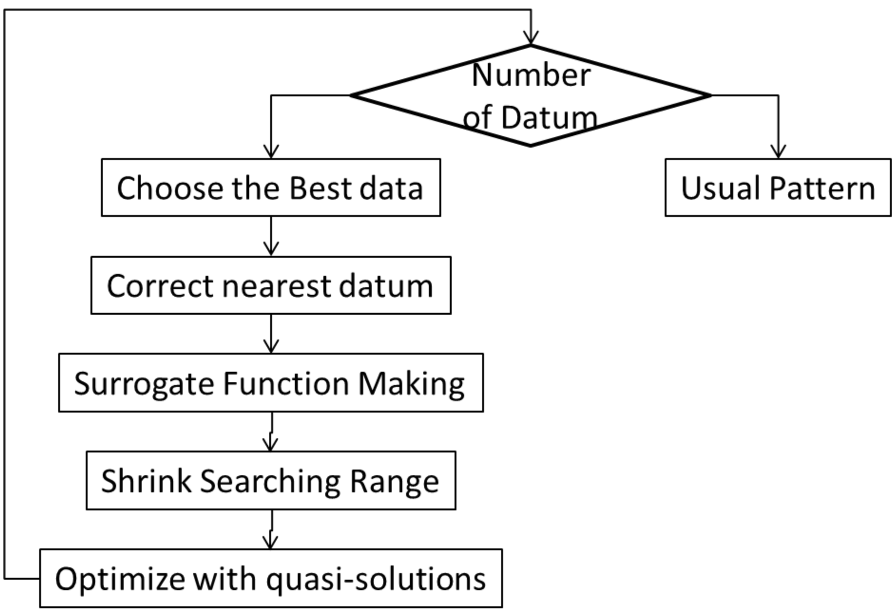


Fig.4 Flow Chart of the proposed method

6. Numerical Example

We use a famous pressure vessel problem as an example. Table 1 shows comparison of the results in some methods.

As this problem is mixed variable problem, so it is very complicated problem. In GRGAs, it needs more than 10,000 function calls to obtain final results, and also it needs more than 6000 function calls to get close to global optima.

Table 1 Comparison of the results of Pressure Vessel Problem

	Sandgren Penalty 1990	Qian GA 1993	Kannan ALM 1994	Lin SA 1992	Hsu GA 1995	Lewis RS+NLP 1996	Arakawa ARGA 1997
R m	1.212	1.481	1.481	N/A	1.316	0.985	0.986
L m	2.990	1.132	1.198	N/A	2.587	5.672	5.626
Ts cm	2.858	2.858	2.858	N/A	2.540	1.905	1.905
Th cm	1.588	1.588	1.588	N/A	1.270	0.953	0.953
g1	0.840	1.000	1.000	N/A	1.000	0.997	1.000
g2	0.747	0.890	0.890	N/A	0.989	0.986	1.000
g3	0.445	0.186	0.182	N/A	0.424	0.938	0.922
g4	1.000	1.000	1.000	N/A	0.831	0.930	1.000
f $	8129.80	7238.83	7198.20	7197.70	7021.67	5980.95	5850.38

Table 2 was the results that we had in the previous study[12] . Correlation means that we have added 10 new data after "Cut 5900", and it is the average of correlation of actual value and RBF outputs for all functions. Indeed, if we add only one optimum solution, we need more than 300~500 function calls to cut 5900 and still keep constraints. Thus, we can safely say that recommendation function and multi-adding points really works better. Table 3 shows the results that we have obtained by using the proposed method. In each case, starting from 50 data, and they are the same compared with the results in Table 2.[13]

Table 2 Results of the previous method

Case	1	2	3
Close to Global	80	80	80
Inside Constraints	100	100	110
Cut 5900	120	120	120
Its Cost	5859.34	5870.29	5865.91
Correlations	0.995	0.994	0.996

We start from the same Initial data with Table 2. However, we add 3 datum for optimization in each case, and add 3 for recommendation function and add 3 for minimization of existence. And we zoomed after we had more than 40 datum at least. Which means, we zoomed after we had more than 80, 120. Table 3 shows the results of the proposed method. In this case, zooming started after we came close to global solutions, so that we could not see the effect of zooming to find global solutions. However, we can still see effectiveness in the number of function calls, and also its accuracy.

Table 3 Results of the proposed method

Case	1	2	3
Close to Global	77	77	77
Inside Constraints	98	98	98
Cut 5900	110	110	122
Its Cost	5854.29	5851.03	5853.92
Correlations	0.996	0.997	0.996

7. Conclusion

In this paper, we propose zooming of searching range for surrogate optimization. As surrogate optimization has a number of datum before we are going to approximate functions. Thus, we can select the best data, and also we can gather nearest datum around the best data. With these datum, we can zoom searching range. It makes approximation of global optima much easier to approximate precisely, and also it makes optimization much easier to find the global solution. In this paper, we have shown the effectiveness of the method through numerical example.

8. References

[1] Kashiwamura, T., Shiratori, M., Yu, K., Optimization of Nonlinear Problem by Using Experimental Design, 1998, Asakura Publishing, In Japanese.

[2] Akira Todoroki, Tetsuya Ishikawa,Design of Experiments for Stacking Sequence Optimizations with Genetic Algorithm using Response Surface Approximation, Composite Structures, 64(3-4),(2004),pp.349-357.
[3] Myers, R.H., Montogomery, D.C., Anderson-Cook, C.M., Response Surface Methodology: Process and Product Optimization Using Designed Experiments,Wiley, 2011.
[4] Arai,H. Suzuki, T, Kaseda, C.Ohyama, K. and Takayama, K., “Bootstrap Re-sampling Technique to Evaluate the Optimal Formulation of Theophylline Tablets Predicted by Non-linear Response Surface Method Incorporating Multivariate Spline Interpolation”, Chem. Pharm, Bull., 55-4,2007,p.586-593.
[5] Simpson, T., Korte, J., Mauery, T. and Mistree, F, Comparison of Response Surface and Kriking Models for Multidisciplinary Optimization, NASA, 1998
[6] Arakawa, M., Nakayama, H., Ishikawa, H., “Optimum Design Using Radial Basis Function Network and Adaptive Range Genetic Algorithms”, Proc Design Engineering Technical Conference'99,(in CD-ROM),ASME,Las Vegas,1999,9
[7] Nakayama, H., Yun, Y., Asada, T., Yoon, Min, “Goal Programming Approaches to Support Vector Machines”, Kowledge-Based Intelligent Information and Engineering Systems, Vol. 2773, (2003), 356-363
[8] Jones, D.R., Schonlau, M., Welch, W.J., Efficient global optimization of expensive black-box functions, Journal of Global Optimization, 13, (1998), 455-492.
[9] Arakawa, M., Nakayama, H., Yun, Y.B., Ishikawa, H., Optimum Design Using Radial Basis Function Networks by Adaptive Range Genetic Algorithms (Determination of Radius in Radial Basis Function Networks),Proceedings of 2000 IEEE International Conference on Industrial Electronics, Control and Instrumentation;21st Century Technology and Industrial Opportunities,in CD-ROM,IEEE,Nagoya,2000.10
[10] Arakawa,M., Nakayama, H., Ishikawa, H., “Approximate Optimization Using RBF Network and Genetic Range Genetic Algorithms: Proposal of Base Function and Basic Consideration”, Trans. JSME, 70-697, 2004, 112-119, (In Japanese)
[11] Arakawa, M. ,Andatsu, A. , Development of Convolute Approximation of Radial Basis Network for Sequential Approximation Optimization, The 7th International Conference on Optimization: Techniques and Application, Kobe, Japan,2007.12
[12] Arakawa, M., Kitayama, S., Scheme for positions of radial basis functions and radius considering supports for accuracy of approximation in convolute RBF , 10th World Congress of Structural and Multidisciplinary Optimization, Orlando, USA, 2013.5
[13] Arakawa, M., Kitayama, S., “Scheme in Setting Position and Radius of RBFin Convolute RBF for Surrogate Optimization”, Proc. Of ENGOPT 2014, Lisbon, 2014.9