

# 유전 알고리즘과 호감도 함수를 이용한 회귀모델의 최적화

안홍락\*(한양대 대학원 정밀기계공학과), 이세현(한양대 기계공학부)

## Optimization of Regression model Using Genetic Algorithm and Desirability Function

H. L. Ahn(Precision Mecha. Eng. Dept. HYU), S. H. Rhee(Precision Mecha. Eng. Dept., HYU)

### ABSTRACT

There are many studies about optimization using genetic algorithm and desirability function. It's very important to find the optimal value of something like response surface or regression model. In this study I indicate the problem using the old type desirability function, and suggest the new type desirability function that can fix the problem better, and simulate the model. Then I'll suggest the form of desirability function to find the optimum value of response surfaces which are made by mean and standard deviation using genetic algorithm and new type desirability function.

**Key Words** : Genetic algorithm (유전 알고리즘), Desirability function (호감도 함수), Response surface (반응 표면), Regression model (회귀 모델)

### 1. 서론

어떠한 반응 표면이나 회귀모델이 주어졌을 경우 그것의 최적값을 찾아내는 것은 매우 중요하다. 본 연구에서는 유전 알고리즘과 새로운 형태의 호감도 함수를 이용하여 평균과 표준편차에 의해 만들어진 두 개의 반응 표면의 최적값을 찾아내는 방법을 제안하였다. 이미 유전 알고리즘과 호감도 함수 접근법을 이용하여 두 반응 표면의 최적값을 찾아내려는 연구가 수행되었으며<sup>(1)</sup>, 이 연구 외에도 호감도 함수를 이용해 최적값을 찾아내기 위한 많은 연구들이 수행되어졌다<sup>(4)(5)(6)(7)</sup>. 본 연구에서는 기존에 사용된 호감도 함수를 사용할 경우에 발생하는 문제를 지적하고 이를 해결하고 보다 좋은 최적화 성능을 얻을 수 있는 호감도 함수의 모델을 제시하였고 이에 대한 시뮬레이션을 수행 하였다.

### 2. 유전 알고리즘

유전 알고리즘은 자연 선택의 원리와 자연계의 생물 유전학에 기본 이론을 둔 병렬적이고 전역적인 알고리즘으로 Darwin 의 적자생존의 이론을 기본 개념으로 한다. 가능한 해들을 정해진 구조의 자료 구조로 표현한 다음 이들을 점차적으로 변형

함으로써 점점 더 좋은 해들을 만들어 내는 유전 알고리즘은 다음과 같은 특성을 가지고 있다.<sup>(2)</sup>

첫째, 유전 알고리즘은 파라미터값(parameter value) 대신에 일정한 길이를 가진 스트링(string)을 사용한다. 일반적인 유전 알고리즘에서는 0 과 1 로 구성된 이진 스트링(binary string)이 주로 이용된다. 둘째, 유전 알고리즘은 탐색 공간에서 일정 크기의 가능해들(possible solutions)을 동시에 고려하여 탐색을 진행한다. 유전 알고리즘의 이러한 병렬 처리는 지역적인(local) 극점으로 수렴하는 것을 방지하고 전역적인(global) 극점을 찾는 데 효과적인 방법이다. 셋째, 유전 알고리즘은 각 스트링의 적합도값(fitness value)만을 사용하기 때문에 그 적합도 함수가 반드시 연속적이거나 미분가능해야 할 필요는 없다. 넷째, 많은 최적화 방법은 결정론적인 전이규칙(deterministic transition rule)을 이용하지만 유전 알고리즘은 확률적인 전이규칙(probabilistic transition rule)을 이용한다.

유전 알고리즘은 크게 재생산(reproduction), 교차(crossover), 돌연변이(mutation)의 과정으로 이루어진다.<sup>(2)</sup>

### 3. 호감도 함수(Desirability function)

호감도 함수 접근법의 기본 개념은 수학적 변환을 통해 여러 개의 응답을 가지는 문제를 하나의 응답을 가지는 문제로 변환 시키는 것이다. 호감도 함수는 다중 응답을 최적화하는 입력 변수를 결정하는 데 유용한 방법으로 이 방법은 계획된 실험과 회귀 분석으로부터 유도된 회귀 모델을 이용한다. 먼저, 각 응답에 대한 호감도 함수를 정의한다. 호감도 함수는 각 응답에 대한 만족도를 나타내는 지수로 만족도가 클수록 1로 접근하고 만족도가 작을수록 0으로 접근하도록 정의한다. 따라서 일반적인 호감도 함수의 범위는 [0,1]이다.

호감도 함수는 최초로 Harrington(1965)에 의해 다음과 같이 제안 되었다.

$$d_j(Y_j(x)) = e^{Y_j(x)} \quad \text{(for the one-sided specification case)} \quad (3.1)$$

$$d_j(Y_j(x)) = e^{-|Y_j(x)|} \quad \text{(for the two-sided case)} \quad (3.2)$$

또한 Derringer and Suich<sup>(4)</sup>는 여러가지 상황에 적용될 수 있는 호감도 함수를 제안하였다

다음으로 위의 식들에 의해 정의된 각 응답에 대한 호감도 함수에 의하여 단일 목적 함수를 만든다. Derringer and Suich는 단일 목적함수로써 각 호감도 함수의 geometric mean 으로 정의된 식 (3.3)와 같은 overall desirability function 을 이용하였다.

$$D(x) = (d_1(Y_1(x)) \cdot d_2(Y_2(x)) \cdots d_m(Y_m(x)))^{1/m} \quad (3.3)$$

여기서  $m$ 은 응답의 개수이다.

마지막으로 overall desirability function 이 정의되면 최적화 알고리즘에 의해 overall desirability function 을 최대화 하는 입력 변수값을 결정하게 된다. Derringer and Suich<sup>(4)</sup>는 최적화 알고리즘으로 direct search method 의 일종인 Hooke-Jeeves search method 를 이용하여 overall desirability function 을 최적화하였다. 이 최적화 방법은 목적함수인 overall desirability function 가 미분 가능해야 할 필요는 없지만 지역적인 최적해를 출력할 가능성이 높다는 단점을 가지고 있다. Castillo, Montgomery, and Mccarville<sup>(5)</sup>는 gradient-based generalized reduced gradient(GRG)알고리즘을 이용하여 overall desirability function 을 최적화 하였다. 이 최적화 방법은 미분 불가능한 점을 포함하는 호감도 함수를 미분 가능한 함수로 변환하는 과정이 필요하고, 시작점에 따라 지역적인 최적값으로 수렴하는 단점을 지니고 있다. 김동철<sup>(1)</sup>은 두 개의 반응 표면을 만들고 이의 overall desirability function 을 유전 알고리즘을 이용하여 최적화 하는 방법을 제안하였다. 이는 호감도

함수의 특별한 변형도 필요 없고, 전역적인 최적해를 빠른 시간 안에 찾는 장점을 가진다. 그러나 앞서 제시한 호감도 함수를 사용하는 경우,  $Y_{minj}, Y_{maxj}$ 의 영역이 좁아질수록 overall desirability function value 0 이 나오는 경우가 많아지게 되고 이로 인하여 최적점에서의 수렴성능이 나빠질 수도 있고, 시뮬레이션 도중 알고리즘 자체에 오류가 발생하는 경우도 생기게 된다.

따라서 변형된 호감도 함수의 여러 형태를 제시하였고 유전 알고리즘과 호감도 함수를 이용해 최적점을 구하였으며 이들의 상호 비교를 통하여 더 좋은 성능을 갖는 호감도 함수를 찾아내었다.

#### 4. 시뮬레이션 방법

본 연구에서 제안된 최적화 방법을 평균과 표준편차에 의해 만들어진 두개의 반응표면에 적용하였다. 두 반응표면은 김동철<sup>(1)</sup>이 제안한 용접 비드 형상 최적화를 위한 반응 표면을 사용 하였고, 두 반응 표면에서 근사적인 최적점의 범위는 [-1,1]이라는 것을 이용하였다. 유전 알고리즘의 각 parameter 는 Table 1 과 같고 유전 알고리즘과 반응표면 분석법을 사용하여 수행한 flowchart 는 Fig. 1 과 같다.

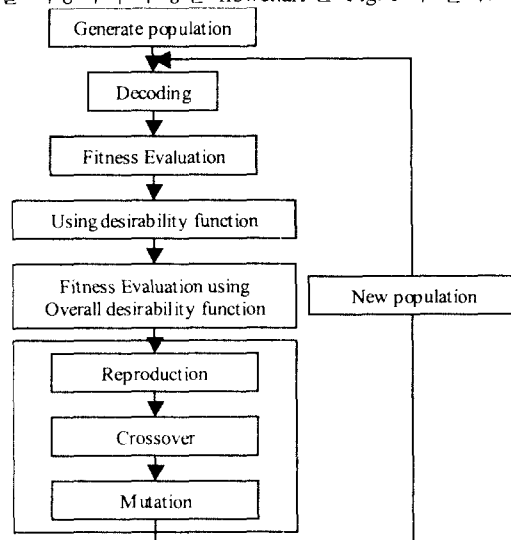


Fig. 1 Flowchart of the whole process

Table 1 Parameter of Genetic Algorithm

Population		30
String	x1	7
	x2	7
Generation		50
Crossover rate		0.95
Mutation rate		0.01

먼저 2 진 스트링을 개체수만큼 발생시켜 이 2 진 스트링을 10 진수로 decoding 한 후 이 10 진수를 fitness function 에 대입하여 각 스트링의 적합도 값을 찾아낸다. 그 후 호감도 함수에 이 적합도 값을 적용하여 각각의 반응 표면에 다른 호감도 함수 값을 구해 내고, overall desirability function value 를 구한다. 이 값들을 구한 후 각각의 스트링이 갖는 적합도에 의하여 이를 재생산하고, 교차, 돌연변이의 과정을 거쳐 새로운 개체를 만들어낸다.

본 연구에 사용된 평균과 표준 편차에 대한 회귀 모델은 다음과 같다.

$$\hat{\mu} = 3.0422 + 0.4967x_1 - 0.4367x_2 + 0.0567x_1^2 - 0.0033x_2^2 \quad (4.1)$$

$$\hat{\sigma} = 0.4813 + 0.1174x_1 + 0.0515x_2 - 0.1515x_1^2 + 0.0115x_2^2 + 0.0462x_1x_2 \quad (4.2)$$

### 5. 시뮬레이션 결과

본 연구에서는 호감도 여러가지 호감도 함수의 형태를 제시하고 그것의 비교를 통하여 가장 바람직한 형태의 호감도 함수를 찾아내고자 하였다. 따라서 기존에 사용된 호감도 함수를 포함하여 5 가지의 호감도 함수 형태를 제시하였고 각각의 비교를 통하여 가장 좋은 경우 호감도 함수를 제시하였다.

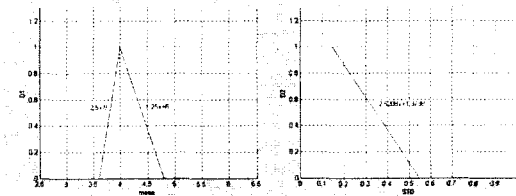


Fig. 2 Desirability function of case 1

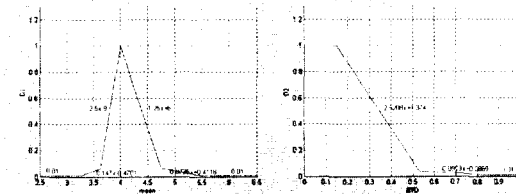


Fig. 3 Desirability function of case 2

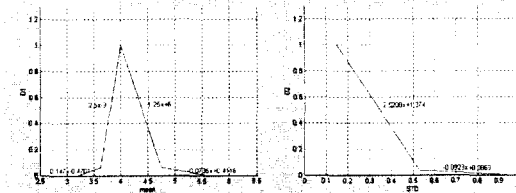


Fig. 4 Desirability function of case 3

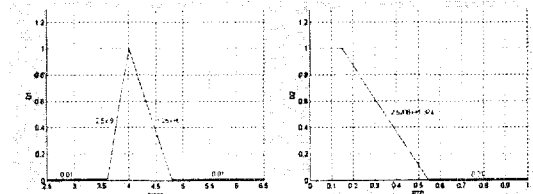


Fig. 5 Desirability function of case 4

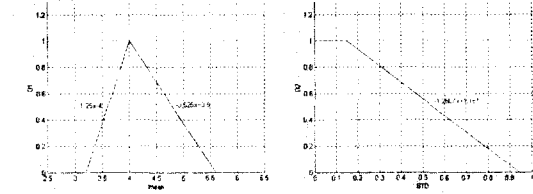


Fig. 6 Desirability function of case 5

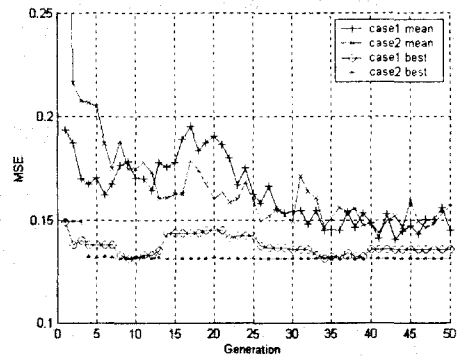


Fig. 7 Comparison result of case 1 and case 2

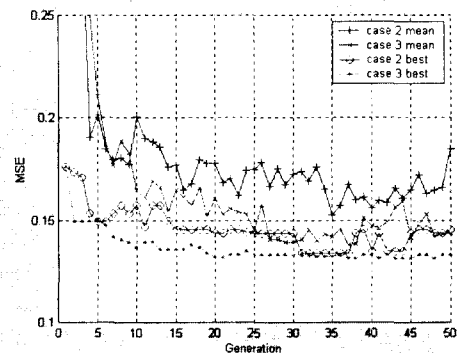


Fig. 8 Comparison result of case 2 and case 3

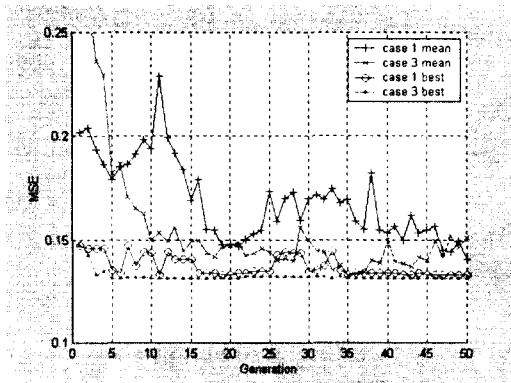


Fig. 9 Comparison result of case 1 and case 3

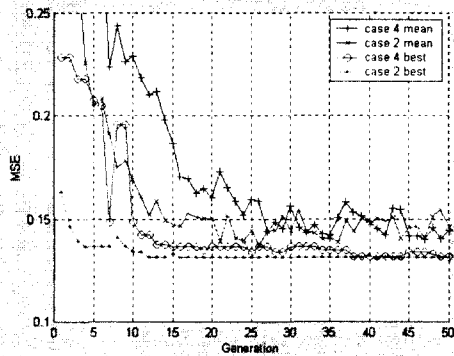


Fig. 10 Comparison result of case 2 and case 4

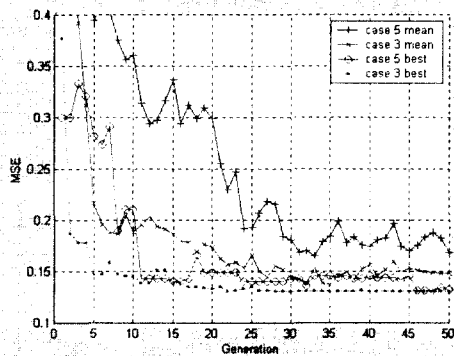


Fig. 11 Comparison result of case 3 and case 5

## 6. 결론

기존의 방법인 case 1 의 호감도 함수를 사용하는 것보다 양단의 호감도 함수값이 0.01 이고 기존 호감도 함수 바깥 부분에 꺾임이 있는 형태의 호감도 함수를 사용하였을 경우 더 좋은 수렴성을 얻을 수 있었으며, 기존의 방법이 overall desirability function 을 모두 0 으로 만들어 알고리즘 오류를 범하는 데 반해 case 2 의 호감도 함수는 이 부분의 오류는 전혀 범하지 않는 장점을 갖는다. 그러나,

양단에 작은 값을 적용해서 알고리즘을 수행하는 경우(case 2) 알고리즘 오류는 전혀 없으나 양단에 작은 값이 없이 꺾인 부분만이 존재 할 경우(case 3)보다 수렴 성능이 떨어짐을 알 수 있다. 또한 네 번째 시뮬레이션 결과에서 알 수 있듯이 꺾인 부분이 있는 경우(case 2)가 꺾임이 없는 경우(case 4)보다 좋은 결과를 나타냄을 알 수 있었고, 마지막 시뮬레이션에서 기존의 호감도 함수를 꺾은 형태가 아닌 넓은 범위에 적용했을 경우(case 5) 꺾인 형태의 호감도 함수를 사용한 경우(case 3)보다 좋지 않은 결과를 얻음을 알 수 있었다.

그러므로 호감도 함수의 폭이 좁아서 생기는 알고리즘의 오류를 해결하기 위해서 호감도 함수의 폭을 넓힐 경우 더 좋은 결과를 얻을 수 있다는 것을 알 수 있고, 기존의 호감도 함수의 양단 밖에서의 호감도 함수 값은 작은 값을 사용하는 것이 좋은 수렴 성능을 얻을 수 있음을 알 수 있다.

## 참고문헌

1. 김동철, "유전 알고리즘과 반응표면 분석법을 이용한 CO<sub>2</sub> 아크 용접 공정의 최적화에 관한 연구", 한양대학교 박사학위 논문, 2000.
2. Goldberg D.E., "Genetic algorithms in search, optimization, and machine learning", Addison-Wesley, 1989.
3. Lawrence Davis, "Handbook of Genetic Algorithm", Van Nostrand Reinhold.
4. Derringer G. and Suich R., "Simultaneous optimization of several response variables", Journal of Quality Technology, Vol. 12, pp. 214-219, 1980.
5. Dell E. Castillo, Montgomery D., and Mccarville D., "Modified Desirability Functions for Multiple Response Optimization", Journal of Quality Technology, Vol. 28, pp. 337-345, 1996.
6. Kim K. and Dennis K. J. Lin, "Dual Response Surface Optimization : A Fuzzy Modeling Approach", Journal of Quality Technology, Vol. 30, pp. 1-10, 1998.
7. Byun J. and Kim K., "A Desirability function approach to the robust design for multiple quality characteristics", Journal of the Korean Institute, Vol. 24, pp. 287-296, 1998.
8. Dell Castillo E. and Montgomery D., "A nonlinear programming solutions to the dual response problem", Journal of Quality Technology Vol. 25, pp. 199-204, 1993.
9. Dennis K. J. Lin and Tu W., "Dual Response Surface Optimization", Journal of Quality Technology, Vol. 27, pp. 34-39, 1995.