

서포트 벡터 회귀를 이용한 블랙-박스 함수의 최적화

곽민정¹⁾, 윤민²⁾

요약

많은 실제적인 공학 설계문제에 있어서, 목적함수의 형태는 설계변수들에 의하여 정확하게 주어지지 않는다. 이러한 환경 하에서, 구조해석, 유체 역학 해석, 열역학 분석과 같은 등과 같은 문제에서 설계변수들의 값이 주어졌을 때 목적함수들의 값은 실제 실험이나 계산상의 실험을 통하여 얻어지게 된다. 일반적으로, 이러한 실험들은 많은 비용이 든다. 이런 경우에는 실험의 횟수를 가능한 적게 하기 위하여, 목적함수의 형태를 예측하는 것과 병행하여 최적화를 수행하게 된다. 반응표면분석(Response Surface Methodology, RSM)은 이러한 접근 방법에서 잘 알려져 있다. 본 논문에서는 목적함수의 예측을 위하여 서포트 벡터 기계(Support Vector Machines, SVM)의 방법을 적용할 것이다. 이러한 접근에서 가장 중요한 과제들 중의 하나는 가능한 실험의 횟수를 적게 하기 위하여 적절하게 표본자료들을 배치하는 것이다. 이러한 목적으로 서포트 벡터의 정보들이 효과적으로 사용되어짐을 보이고 제안한 방법의 효율성은 공학 설계문제에서 잘 알려진 수치 예제를 통하여 보인다.

주요용어: 서포트 벡터 회귀; 유전자 알고리즘; 대역정보와 국소정보; 최적화.

1. 서론

많은 실제적인 공학문제에 있어서, 블랙-박스 목적함수의 형태들은 설계변수에 의하여 정확하게 알려져 있지 않으나, 이러한 함수 값들은 실제적으로나 계산 실험들을 통하여 얻어진 값들로써 얻어지게 된다. 이러한 경우에 있어서, 구조분석, 유체역학 해석, 열역학 해석 등과 같은 여러 분야에서 목적함수들을 평가하기 위하여 적용된다. 종종 이러한 분석들은 계산 시간이 현저하게 소모된다. 그러므로, 기존의 방법들에 의하여 이러한 함수들이 최적화 되어진다면, 해를 얻기 위하여 비현실

1) (450-701) 경기도 평택시 용이동 111번지, 평택대학교 전산통계학과, 교수.

E-mail: mjkwak@ptuniv.ac.kr

2) (143-701) 서울시 광진구 화양동 1번지, 건국대학교 응용통계학과, 강의교수.

Correspondence: myoon515@konkuk.ac.kr

적으로 많은 시간이 요구될 것이다. 이러한 상황에는 필요한 실험점들을 가능한 적게 해야한다.

이러한 종류의 문제들은 공학에서와 마찬가지로 지리학에서도 연구되어 왔고, 목적함수를 최적화하는 효과적인 실험점들을 찾는 방법을 개발하는 것이 본 논문의 목적이다. 가능하면 적은 수의 실험점들에 의하여 잘 근사된 블랙-박스 목적함수를 찾는 것이 중요한 목적 중의 하나이다.

반응표면분석은 본 논문의 목적에 가장 적합할 것이다. 종종 반응표면 모형은 광범위한 방법들을 포괄하는 이름으로 사용된다. 무엇보다도 실험계획에서 사용되는 방법들은 유명하다. 그러나, 그러한 방법들 중에서 많은 것들이 단지 설계변수 공간의 통계적 분석에 근거하여 실험점들을 선택한다. 그것들은 완만한 비선형성을 가지는 블랙-박스 함수에서 좋은 근사함수를 제공한다. 그러나, 명백하게 블랙-박스 함수가 심한 비선형성을 가지는 경우라면, 설계변수 공간상에서의 통계적인 성질뿐만 아니라 함수의 형태를 동시에 고려하는 방법들이 더 좋은 수행능력을 얻을 수 있을 것이다.

최근에 여러 학자들에 의하여 블랙-박스 함수를 근사하기 위하여 Radial Basis Function 네트워크(RBFN)의 적용을 제안해 왔다 (Haykin, 1999; Jones 등, 1998; Sacks 등, 1989; Schonlau, 1997). 따라서 실험 점들은 블랙-박스 함수의 대역적 정보와 국소 정보를 동시에 고려하는 것으로 주어진다. 본 논문은 서포트 벡터들이 이러한 목적에 효과적으로 부합됨을 보일 것이다.

2. 서포트 벡터 회귀식의 고찰

SVM의 주요 아이디어는 통계적 학습이론에 기초한 패턴분류 문제들에 대하여 커널을 이용한 최대 여유도(margin)를 갖는 선형 분류기를 얻는 것이다. SVM은 회귀 문제에 적용하는 것들이 제안되어 왔다. 입력공간 X 에서 주어진 표본 자료들을 (\mathbf{x}_i, y_i) 로 나타내자. 여기서 $i = 1, \dots, \ell$ 이다. X 상에서의 회귀 함수를

$$f(\mathbf{x}) = \langle \boldsymbol{\beta}, \mathbf{x} \rangle + \beta_0$$

와 같이 표현된다고 가정하자. 선형 ε 손실함수는

$$L^\varepsilon(\mathbf{x}, y, f) = |y - f(\mathbf{x})|_\varepsilon = \max(0, |y - f(\mathbf{x})| - \varepsilon)$$

와 같이 정의 된다. 유사하게 이차 ε 손실함수는

$$L_2^\varepsilon(\mathbf{x}, y, f) = |y - f(\mathbf{x})|_\varepsilon^2$$

로 정의된다. 손실함수들은 둘 다 실제값의 어떤 거리 내부의 오차들은 무시한다. 본 논문에서는, 이차 ε 손실함수를 사용한다. 이차 ε 손실함수의 합을 최소화함으로서 아래와 같이

$$\text{minimize} \quad \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^{\ell} L_2^{\varepsilon}(\mathbf{x}_i, y_i, f)$$

회귀함수식을 최적화할 것이다. 여기서 C 는 복잡성과 손실 사이의 트레이드-오프를 나타내는 측도이다. 일반적으로 비선형인 경우에 대한 회귀를 적용하기 위하여, X 상에서 주어진 실험점 (\mathbf{x}_i, y_i) $i = 1, \dots, \ell$ 는 비선형 변환함수 ϕ 에 의하여 고차원의 특징공간(feature space)으로 변환한다. 이제 우리의 회귀문제는 아래의 문제를

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^{\ell} (\xi_i^2 + \xi'_i) \\ \text{subject to} \quad & \langle \beta, \phi(\mathbf{x}_i) \rangle + \beta_0 - y_i \leq \varepsilon + \xi_i, \\ & y_i - \langle \beta, \phi(\mathbf{x}_i) \rangle - \beta_0 \leq \varepsilon + \xi'_i, \\ & \xi_i, \xi'_i \geq 0, \quad i = 1, \dots, \ell \end{aligned} \tag{2.1}$$

최적화하는 문제로 축소할 수 있다. 여기서 slack 변수 ξ_i 는 대상(target)의 값에서 ε 보다 초과한 양을 나타내고, 다른 변수 ξ'_i 는 대상값에서 ε 보다 부족한 부분의 양을 나타낸다. 위의 수식에 대한 이차계획문제의 최소화는 라그랑주 승수 α, α' 를 도입한 쌍대문제의 최대화와는 동치이다. (2.1)에 대한 쌍대문제에서

$$\phi(\mathbf{x}_i)^t \phi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j) \tag{2.2}$$

를 만족하는 커널함수 $K(\mathbf{x}_i, \mathbf{x}_j)$ 를 사용하여 최종적으로 아래의 공식을 얻는다.

$$\begin{aligned} \text{maximize} \quad & -\frac{1}{2} \sum_{i,j=1}^{\ell} (\alpha'_i - \alpha_i)(\alpha'_j - \alpha_j) \left\{ K(\mathbf{x}_i, \mathbf{x}_j) + \frac{1}{C} \delta_{ij} \right\} \\ & + \sum_{i=1}^{\ell} y_i (\alpha'_i - \alpha_i) - \varepsilon \sum_{i=1}^{\ell} (\alpha'_i + \alpha_i) \\ \text{subject to} \quad & \sum_{i=1}^{\ell} (\alpha'_i - \alpha_i) = 0, \quad \alpha'_i, \alpha_i \geq 0, \quad i, j = 1, \dots, \ell. \end{aligned} \tag{2.3}$$

여기서 δ_{ij} 는 $i = j$ 인 경우는 1이고, 그 외의 경우는 0인 Kronecker 델타이다. Kaurush-Khun-Tucker의 보충조건으로

$$\begin{aligned}\alpha_i(f(\mathbf{x}_i) - y_i - \varepsilon - \xi_i) &= 0, \\ \alpha'_i(y_i - f(\mathbf{x}_i) - \varepsilon - \xi'_i) &= 0, \\ \xi_i \xi'_i &= 0, \quad \alpha_i \alpha'_i = 0, \quad i = 1, \dots, \ell\end{aligned}$$

를 얻는다. 일차계획문제 (2.1) 보다는 그 쌍대문제인 (2.3)에서 가우지안 커널

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{r^2}\right) \quad (2.4)$$

를 적용하여 문제를 풀게된다. RBFN에서와 마찬가지로, 폭 r 은 $d_{\max}/\sqrt[n]{nm}$ 에 의해서 잘 추정이 될 것이고 (Haykin, 1999), 여기서 d_{\max} 는 자료점들 사이의 최대의 거리를 나타낸다. 또한 n 은 X 에 있어서 실험점들의 차원이고, m 은 실험점들의 개수이다. 고차원의 특징공간에서 커널을 사용한 회귀함수식은

$$f(\mathbf{x}) = \sum_{i,j=1}^{\ell} \alpha_i^* K(\mathbf{x}_i, \mathbf{x}_j) + \beta_0^*, \quad \alpha_i^* = \alpha'_i - \alpha_i \quad (2.5)$$

와 같이 주어진다. 만약 $\alpha_i^* \neq 0$ 이면 i 번째 실험점은 추정된 함수 $f(\mathbf{x})$ 들레의 $\pm \varepsilon$ 의 띠 상에 있는 서포트 벡터이다. 서포트 벡터들에 의하여 회귀함수 $f(\mathbf{x})$ 가 표현되기 때문에 이러한 서포트 벡터들은 중요한 실험점들로 간주할 수 있다.

3. 블랙-박스 목적함수들의 최적화

SVM에 의하여 블랙-박스 목적함수들을 추정한 후에, 이것들을 최적화 하기 위하여 유전자 알고리즘을 사용한다. 이 유전자 알고리즘은 사용하기에 간단하여 적용하기가 쉽다. 또한 대역적 최적해의 근사해를 제공하기 때문이다. 연속형 설계 변수들의 경우에, 단순한 방법으로 연속형 설계변수들을 처리할 수 있는 BLX- α 알고리즘을 적용한다. 설계변수들의 연속적인 값들은 개체들의 코드들로서 이용한다. 자녀들은 그들의 부모들이 포함된 직사각형의 공간내부에서 무작위로 생성된다. 연속형의 설계변수를 갖는 문제들에서 돌연변이가 없는 경우에는 BLX- α 방법이 효과적이라고 알려져 있다 (Eshelman 와 Schaffer, 1993).

제안한 방법의 알고리즘은 아래와 같이 요약할 수 있다.

- 1 단계: SVM으로 목적함수의 형태를 추정한다.

2 단계: 유전자 알고리즘으로 추정된 목적함수에 대하여 최적점을 추정한다.

3 단계: S_0 까지에서 나타난 최적점들의 갯수($C_{\#}$)를 센다. 현재 최적점을 중심으로하는 고정된 변의 길이 ℓ 을 가지는 정사각형에 의해 주어지는 S_0 근방(neighborhood)이라 하자.

4 단계: 만일 C_x 가 사전에 주어진 C_x^0 보다 같거나 혹은 큰 경우, 또는 반복을 종료한다.

4. 추가 실험점들의 선택방법

만일 현재의 해가 만족스럽지 못하다면, 즉 정지조건을 만족하지 못하면, 블랙-박스 목적함수의 근사를 개선하기 위하여 추가적인 실험점들이 필요하다. 현재의 최적 실험점이 추가적인 자료에서 선택이 되면 추정된 최적점들은 국소 최대(최소)점에 수렴할 것이다. 이러한 사실은 목적함수를 예측하는데 있어서 대역적 정보의 결여에 기인한다. 다른 한 편, 기존의 자료점들에서 멀리 떨어진 점들을 추가 실험점으로 선택하면 최적점 근방에서 좀 더 세밀한 정보를 얻기가 어려울 것이다. 따라서 높은 정밀도를 갖는 해를 얻기가 어려울 것이다. 이것은 최적점 근방에서 불충분한 정보 때문이다.

블랙-박스 목적함수의 대역적 정보와 국소적 정보를 제공하는 균형잡힌 표본을 얻는 것이 중요하다. Nakayama 등 (2002)은 목적함수를 추정하는데 대한 대역적 정보와 최적점 근방의 국소적 정보를 동시에 고려하는 방법을 제안하였다. 두 종류의 추가적인 실험점들은 목적함수 형태의 재학습에 대해 동시에 택할 수 있다. 그들을 중에 하나는 추정된 최적점 근방의 국소적 정보를 추가하기 위하여 현재의 최적점 근방에서 선택한다. 이 근방의 크기는 수렴하는 과정 동안에 조절된다. 대역적 정보를 위한 추가점은 목적함수의 형태를 더 잘 예측하기 위하여 현재 최적점에서 떨어진 실험점을 선택한다. 전자의 추가점은 현재 최적점 근방에서 좀 더 세밀한 정보를 제공한다. 후자의 실험점은 국소 최대(최소)점으로 수렴하는 것을 막아 준다. 현재 최적점의 근방은 정사각형 S 에 의하여 주어지는데 이 정사각형의 중심은 현재의 최적점이고 변의 길이는 ℓ 이다. 중심은 현재 최적점이고 한 변의 길이는 ℓ_0 인 정사각형을 S_0 라 하자. S 는 과거의 S_0 에서 연속적으로 나타난 최적점의 수 C_x 에 따라서 축소된다. 즉,

$$\ell = \ell_0 \times \frac{1}{C_x + 1}$$

와 같다. 첫번째 추가점들은 정사각형 S 내부에서 무작위로 선택된다. 두번째 추가 실험점은 S 의 외곽에 있는 기존의 학습자료들이 드물게 있는 부분에서 선택된다.

어떤 정지조건들을 만족할 때까지, 목적함수들의 근사치를 개선하기 위하여 몇몇 개의 실험점들을 추가한다. 각각의 반복 단계에서 대역적 정보와 국소적 정보에 대한 추가를 위하여 두 종류의 실험점들을 추가한다. 첫 번째는 현재의 최적점의 근방에서 주어진 국소적 정보에 대하여 추가 실험 자료들을 선택한다. 이것을 하기 위하여, 첫번째 추가되는 실험 자료들의 어떤 후보자료는 정방형의 S_0 에서 무작위로 생성된다. 주어진 대역적 정보에 대한 두 번째의 추가 실험 자료들은 기존의 실험점들이 거의 없는(sparse) 부분에서 선택한다. 두 번째 추가 실험 자료들의 어떤 후보점들은 S_0 바깥의 임의의 부분에서 생성된다. 주어지는 국소적 정보와 대역적 정보에 대한 후보점들 간의 차이는 S_0 내부에서 생성된 자료인지 혹은 외부에서 생성된 자료인지에 있다. 기존에 존재하는 실험점들의 밀도를 고려하여 후보의 각각의 집합들 중에서 후보점들의 하나를 선택한다. 즉, 기존의 실험점들이 거의 없는(sparse) 부분에서 실험점을 찾는다.

자료들을 추가하는 순서는 아래와 같다.

- 1 단계: 실험점 x_i^C 의 여러 후보점들을, 예를 들어 $i = 1, \dots, \ell$, 무작위로 생성한다.
- 2 단계: 각각의 후보 실험점과 모든 기존의 실험점들 사이의 거리를 측정한다.
- 3 단계: 가장 근접한 자료점에서 최대거리를 가지는 후보 실험점들 중에서 가장 작은 개수를 선택한다.
- 4 단계: 3 단계에서 선택된 실험점들의 각 후보점들에서 α_C^* 의 민감도를 고려한다.
- 5 단계: 새로운 실험점으로서 $|\alpha_C^*|$ 의 값이 가장 큰 자료를 선택한다.

$|\alpha_C^*|$ 의 값은 아래와 같이 계산된다. 우선 새로운 자료의 후보점을 x^C 라 하자. $y(x^C)$ 의 실제값을 알 수 없기 때문에, 대신에 추정된 $f(x^C)$ 의 값을 사용한다. x^C 에서 α_C^* 의 민감도를 계산하기 위하여, 1차 계획문제 (2.1)에서 아래의 제약조건

$$\begin{aligned} \sum_{i=1}^{\ell+1} \beta_i \phi(\mathbf{x}_i) - \beta_0 &\leq \varepsilon + \xi_{\ell+1}, \\ f(\mathbf{x}^C) - \sum_{i=1}^{\ell+1} \beta_i \phi(\mathbf{x}_i) - \beta_0 &\leq \varepsilon + \xi'_{\ell+1}, \\ \xi_{\ell+1}, \quad \xi'_{\ell+1} &\geq 0 \end{aligned} \tag{4.1}$$

(4.1)식을 추가한다.

새로운 제약식 (4.1)을 추가한 후에, 특징공간에서 쌍대문제는

$$\begin{aligned} \text{maximize} \quad & -\frac{1}{2} \sum_{i,j=1}^{\ell+1} \alpha_i^* \alpha_j^* \left\{ K(\mathbf{x}_i, \mathbf{x}_j) + \frac{1}{C} \delta_{ij} \right\} + \sum_{i=1}^{\ell} y_i \alpha_i^* \\ & - \varepsilon \sum_{i=1}^{\ell+1} \alpha_i^* + f(\mathbf{x}^C) \alpha_{\ell+1}^* \\ \text{subject to} \quad & \sum_{i=1}^{\ell+1} \alpha_i^* = 0, \quad \alpha_i, \alpha'_i \leq 0 \quad i = 1, \dots, \ell+1 \end{aligned} \quad (4.2)$$

와 같이 주어진다. 여기서 $\alpha_i^* = \alpha'_i - \alpha_i$ 이다. (4.2)의 해에서 α_C^* 의 가장 큰값을 가지는 새로운 자료 x^C 를 선택하여 사용한다.

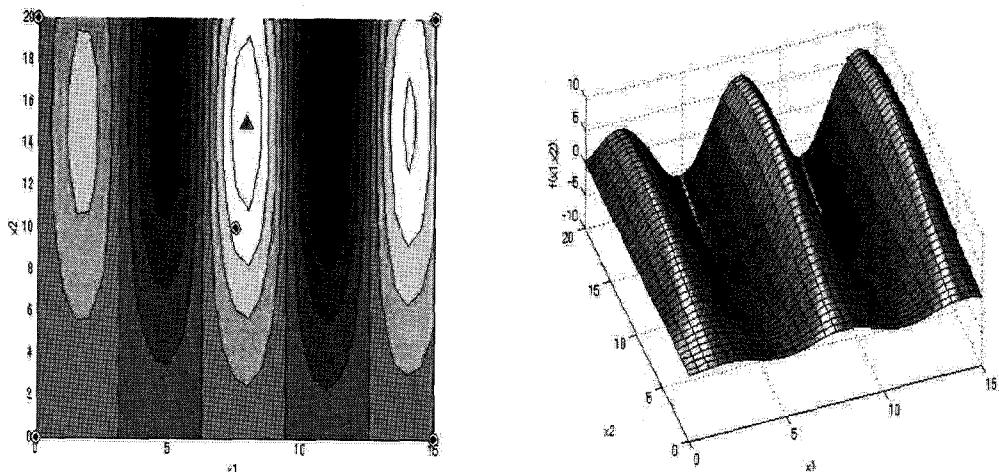
5. 예제

본 논문에서 사용된 예제는 공학설계 분야에서 매우 잘 알려진 예제로서 제안한 방법의 우수성을 보이고자 사용하였다. 우선, 아래에 주어진 함수를 최대화하는 예제를 고려하자.

$$\begin{aligned} f(\mathbf{x}) = & 10 \exp\{-0.01(x_1 - 10)^2 - 0.01(x_2 - 15)^2\} \sin(x_1) \\ & 0 \leq x_1 \leq 15, \quad 0 \leq x_2 \leq 20. \end{aligned} \quad (5.1)$$

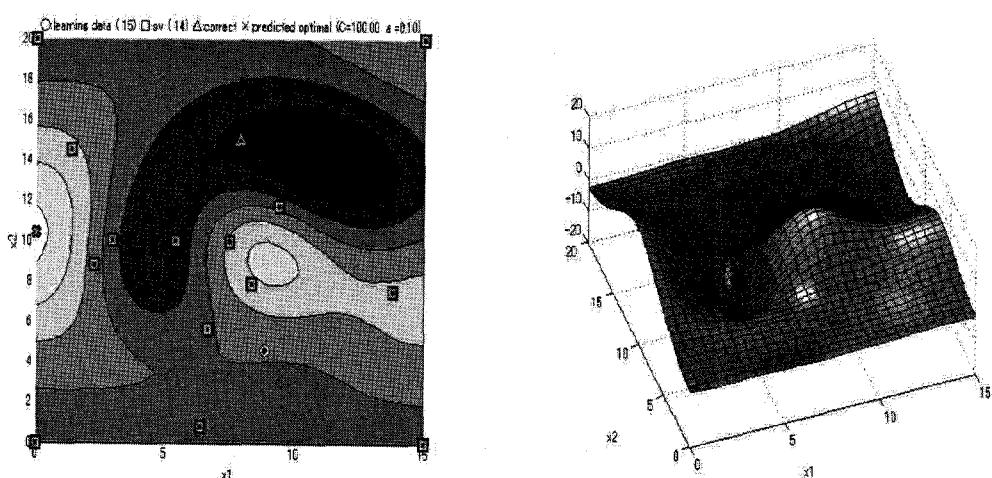
이 함수는 $(x_1, x_2) = (7.8960, 15.0000)$ 에서 최대값 $f = 9.5585$ 를 가진다. 초기 실험점들로 \odot 로 표시한 $(x_1, x_2) = (0, 0), (15, 0), (0, 20), (7.5, 10), (15, 20)$ 의 5개의 점들을 선택하였고, 정확한 최적점은 \triangle 으로 나타내었고, 예측된 최적점은 \times 로 나타내었다. 서포트 벡터 기계에서 $C = 100.0$ 과 $\varepsilon = 0.1$ 로 설정하고, 유전자 알고리즘에서 $\ell_0 = 4.0$, $C_0 = 20$, 부모집단은 10개로, 그리고 자녀집단은 200으로 두고 실험을 하였다. 모의실험의 결과는 아래의 표 5.1에서 나타내었다.

모의실험에서 추정된 함수의 중간과정은 그림 5.2와 그림 5.3에서 볼 수 있다. 최종단계의 그림은 그림 5.4에서 나타내었다. 비교의 편리를 위하여, 유전자 알고리즘만을 사용한 결과를 표 5.2에서 나타내었다. 제안된 방법과 동일하게 부모집단을 10으로 두고 실험을 하였다. 주어진 고정된 범위내에서 부모들과 자식들간에 높은 횡적관계를 가지는 상위 10개의 개체들이 수렴을 하는 경우에 계산을 멈추었다. 표 5.2는 계산이 멈출 때까지 생성된 개체(# 분석으로 나타내어진)들의 전체 갯수를 나타내었다. * 표시는 수렴하는 국소 최대값을 나타낸다. 서포트 벡터들의 갯수는 평균 47.6개였고, 즉, 만일 추가되는 실험점들이 가장 효과적인 것으로 47.6개 까지 분석의 개수를 감소시킬 것으로 기대할 수 있다.



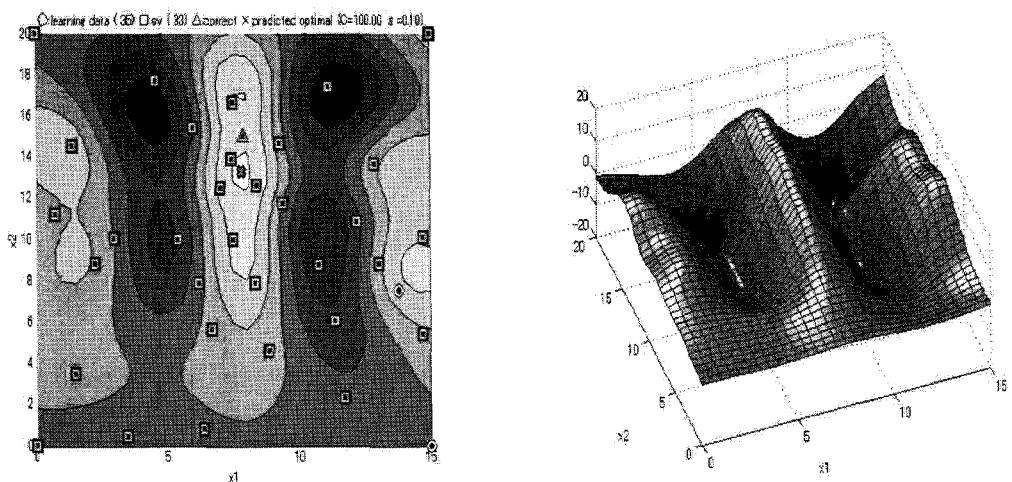
$$f(x_1, x_2) = f(7.8960, 15.0000) = 9.5585$$

그림 5.1: 예제



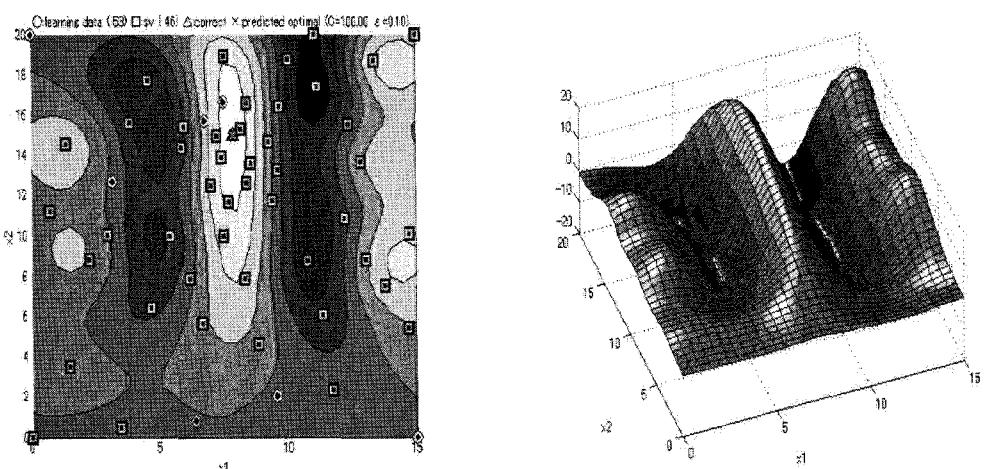
15개의 표본자료와 14개의 서포트 벡터들

그림 5.2: 15개의 표본자료들을 이용한 결과



35개의 표본자료와 33개의 서포트 벡터들

그림 5.3: 35개의 표본자료들을 이용한 결과



53개의 표본자료와 46개의 서포트 벡터들

그림 5.4: 53개의 표본자료들을 이용한 결과

표 5.1: 제안된 방법에 의한 모의실험 결과

	분석의 수	sv의 갯수	x_1	x_2	$f(x_1, x_2)$
1	49	44	7.879	15.465	9.5366
2	55	51	7.846	15.666	9.5044
3	53	50	7.892	14.465	9.5312
4	53	46	7.896	15.022	9.5585
5	61	47	7.897	14.724	9.5515
ave	54.2	47.6	7.882	15.069	9.5364
σ^2			0.0006	0.2045	0.0008

표 5.2: 유전자 알고리즘만을 이용한 모의실험 결과

	분석의 수	x_1	x_2	$f(x_1, x_2)$
1	100	8.008	14.678	9.4869
2	60	7.980	13.635	9.3532
3	100*	14.093	13.851	8.3383
4	90	8.000	12.650	8.9946
5	80	7.902	14.937	9.5580
ave	86.0	9.1966	13.954	9.1462
σ^2		7.6876	1.7530	0.3708

6. 결론

설계변수에 의하여 목적함수의 형태가 정확하게 알려지지 않은 경우에, 효과적으로 새로운 실험점들을 어떻게 추가를 할 것인가는 매우 중요한 주제이다. 제안한 방법은 새로운 실험점들을 효과적으로 제공하는 서포트 벡터 기계에 의한 각각의 후보 실험점들의 민감도를 고려한 방법으로서 국소적 정보와 대역적 정보를 동시에 고려하여 실험점을 선택하는 방법이라는 것을 보였다. 본 논문에서는 제안한 방법으로 몇몇의 실제문제들에 적용을 해보았다. 이 방법은 실제적인 적용의 관점에서 효과적이고 응용가능성이 높다고 기대된다.

참고문헌

Cristianini, N. and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.

- Eshelman, L. J. and Schaffer, J. D. (1993). Real-coded genetic algorithms and interval-schemata. In *Foundations of Genetic Algorithms 2* (Whitley, L. D., ed.), Morgan Kaufman.
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. 2nd ed., Prentice Hall.
- Hsu, Y. H., Sun, T. L. and Leu, L. H. (2000) A two-stage sequential approximation method for nonlinear discrete variable optimization, *ASME Design Engineering Technical Conference*, Boston, 197–202.
- Jones, D. R., Schonlau, M. and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, **13**, 455–492.
- Kannan, B. K. and Kramer, S. N. (1994). An augmented lagrange multiplier based method for mixed integer discrete continuous optimization and its applications to mechanical design. *Journal of Mechanical Design*, **116**, 405–411.
- Myers, R. H. and Montgomery, D. C. (1995). *Response Surface Methodology: Process and Product Optimization using Designed Experiments*. John Wiley & Sons, New York.
- Nakayama, H., Arakawa, M. and Sasaki, K. (2002). Simulation based optimization for unknown objective functions. *Optimization and Engineering*, **3**, 201–214.
- Radcliffe, N. J. (1991). Formal analysis and random respectful recombination. In *Proceedings of the Forth International Conference on Genetic Algorithms*, 222–229.
- Sacks, J., Welch, W. J., Mitchell, T. J. and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical Science*, **4**, 409–435.
- Schonlau, M. (1997). Computer Experiments and Global Optimization. PhD. thesis, University of Waterloo, Ontario, Canada.
- Schölkopf, B. and Smola, A. J. (2002), *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. The MIT Press.
- Zabinsky, Z. B. (1998). Stochastic methods for practical global optimization. *Journal of Global Optimization*, **13**, 433–444.
- Zhang, C. and Wang, H. P. (1993). Mixed-Discrete nonlinear optimization with simulated annealing. *Engineering Optimization*, **21**, 277–291.

[2007년 12월 접수, 2008년 1월 채택]

Using Support Vector Regression for Optimization of Black-box Objective Functions

Minjung Kwak¹⁾, Min Yoon²⁾

Abstract

In many practical engineering design problems, the form of objective functions is not given explicitly in terms of design variables. Given the value of design variables, under this circumstance, the value of objective functions is obtained by real/computational experiments such as structural analysis, fluidmechanic analysis, thermodynamic analysis, and so on. These experiments are, in general, considerably expensive. In order to make the number of these experiments as few as possible, optimization is performed in parallel with predicting the form of objective functions. Response Surface Methods (RSM) are well known along this approach. This paper suggests to apply Support Vector Machines (SVM) for predicting the objective functions. One of most important tasks in this approach is to allocate sample data moderately in order to make the number of experiments as small as possible. It will be shown that the information of support vector can be used effectively to this aim. The effectiveness of our suggested method will be shown through numerical example which is well known in design of engineering.

Keywords: Support vector regression; genetic algorithm; global and local information; optimization.

1) Associate Professor, Department of Computer Science & Statistics, Pusan National University,
111 Yonggi-Dong, Pusan, Kyungki-Do 450-701, Korea. E-mail: mjkwak@pusan.ac.kr
2) Teaching Professor, Department of Applied Statistics, Konkuk University, 1 Hwayang-
Dong, Seoul 143-701, Korea. Correspondence: myoon515@konkuk.ac.kr