서포트 벡터 회귀를 이용한 제어기 설계

Design of controller using Support Vector Regression

황 지 환*, 곽 환 주**, 박 귀 태***
(Ji-Hwan Hwang*, Hwan-Joo Kwak**, Gwi-Tae Park***)

Abstract - Support vector learning attracts great interests in the areas of pattern classification, function approximation, and abnormality detection. In this pater, we design the controller using support vector regression which has good properties in comparison with multi-layer perceptron or radial basis function. The applicability of the presented method is illustrated via an example simulation.

Key Words : Support vector machine(regression), multi-layer perceptron, cart-ball controller

1. 서 론

서포트 벡터 학습 방법은 이론과 각종 응용 사례가 보고되면서, 지능시스템 분야에서 매우 주요한 도구로 되고 있다 [1]. 서포트 벡터 학습 방법은 최대 마진을 이용하여 유도할 수 있는데 일반화 능력(generalization capability)을 효과적으로 보장할 수 있어 패턴 분류[2], 함수 근사[3], 비정상 상태 탐지[4] 등의 여러 가지 문제에 다양하게 적용되었다.

서포트 학습방법은 한 개의 은닉층을 갖는 신경망(multi-layer perceptron) 또는 RBFN(radial basis function network)에 비해 다음과 같은 장점을 가지고 있다. 첫째, 은닉층의 은닉노드 수를 자동으로 결정할 수 있다. 둘째, 기울기 강화법 등의 학습방법이 가지는 지역적 최적해(local optimum)로 수렴하는 문제가 없어져 반드시 전역적 최적해(global optimum)를 찾을 수 있다. 셋째, 통계적 학습이론(statistical learning theory)으로 설명될 수 있기 때문에 일반화 능력이 우수한 결과를 학습 결과로 얻을 수 있다. 본 논문에서는 이러한 서포터 벡터의 응용 사례 중 함수 근사 기능, 즉 서포터 벡터 회귀를 이용하여 시스템의 제어기를 설계하고 자 한다. 주어진 데이터에서 일정한 성능지수(performance index)를 최적화할 수 있는 근사함수(approximating function)를 찾아서 그 결과값이 제어입력이 되도록 한다.

2. 서포트 벡터 회귀

서포트 벡터 학습방법은 원래 패턴 분류를 위해 Vapnik과 공동연구자들에 의해 개발되었고, 이후 서포터 벡터 회귀로 발전되었다. 본 논문에서는 입실론(epsilon) 서포터 벡터 회귀 방법을 이용하여 함수 근사를 한다. 입력공간 X에서 주어진 학습데이터를 (x_i, y_i)로 나타내자. 여기서 $i=1,.....,l$ 이다. 주어진 학습데이터에 대해 입실론-무반응 오차(ϵ-insensitive error)의 총합을 작게 하는 매끄러운 근사함수를 찾는 것이 서포터벡터 회귀의 목적이다. X 공간상에서 회귀함수를 아래와 같이 나타낸다.

$$f(x) = \langle \beta, x \rangle + \beta_0 \quad (1)$$

입실론-무반응 오차함수는

$$L^{\epsilon}(x, y, f) = |y - f(x)|_{\epsilon} = \max(0, |y - f(x) - \epsilon|) \quad (2)$$

와 같이 정의된다. 서포트 벡터 회귀는 입실론-무반응 오차함수의 합을 최소화함으로써 아래와 같이 회귀 함수식을 최적화 한다.

$$\min \frac{1}{2} \| \beta \|^2 + C \sum_{i=1}^{l} L^{\epsilon}(x_i, y_i, f) \quad (3)$$

여기서 C는 복잡성과 손실 사이의 trade-off를 나타내는 일반화 모수(generalization parameter, regularization parameter)이다. 일반적으로 비선형인 경우에는 X상의 입력공간(input space)에서 비선형 변환함수 Φ에 의하여 고차원의 특징공간(feature space)으로 변환한다.

여기에, 여러 분야에서 널리 사용되고 있는 가우시안 RBF 커널을 이용한 특징공간으로의 변환함수 Φ를 사용하면 다음과 같은 형태의 커널 트릭이 성립된다[1].

$$< \Phi(x), \Phi(y) > = k(x, y) = \exp(- \| x - y \|^2)/2\sigma^2) \quad (4)$$

저자 소개
* 正 會 員 : 高麗大學 電子電氣工學科 博士課程
** 準 會 員 : 高麗大學 電子電氣工學科 碩博士統合課程
*** 正 會 員 : 高麗大學 電氣電子電波工學科 正敎授・工博

이제 회귀문제는 아래의 문제를 최적화하는 문제로 귀결된다.

$$\min \frac{1}{2}\|\beta\|^2 + C\sum_{i=1}^{l}(\xi_i+\xi_i^*) \tag{5}$$
$$\text{subject to } \langle\beta, \Phi(x_i)\rangle+\beta_0-y_i \le \epsilon+\xi_i,$$
$$y_i-\langle\beta, \Phi(x_i)\rangle-\beta_0 \le \epsilon+\xi^{*}_i,$$
$$\xi_i, \xi_i^* \ge 0, \qquad i=1,\dots,l$$

여기서 slack 변수 ξ_i, ξ_i^*는 ϵ보다 초과한 값과 부족한 값을 각각 나타낸다. 이러한 최적화 문제에 대응하는 라그랑제 함수(Lagrange function)는 다음과 같다.

$$L=\frac{1}{2}\|\beta\|^2+C\sum_{i=1}^{l}(\xi_i+\xi_i^*)-\sum_{i=1}^{l}\alpha_i(\epsilon+\xi_i-y_i+<\beta, \Phi(x_i)>+\beta_0)$$
$$-\sum_{i=1}^{l}\alpha^{*}_i(\epsilon+\xi^{*}_i+y_i-<\beta, \Phi(x_i)>-\beta_0)$$
$$-\sum_{i=1}^{l}(\eta_i\xi_i+\eta^{*}_i\xi^{*}_i),\ \text{단},\ \alpha^{(*)}_i, \eta^{(*)}_i \ge 0 \tag{6}$$

여기에서 $\beta, \beta_0, \xi_i, \xi^{*}_i$는 주변수(primal variables) 이고, $\alpha_i, \alpha_i^*, \eta_i, \eta_i^*$는 라그랑제 승수(Lagrange multipliers)로 도입된 쌍대 변수(dual variable)이다. 식(5)의 최소값을 찾는 문제는 라그랑제 승수로 도입된 쌍대 문제의 최대화와 동치이다. 또한 비선형 회귀문제를 위한 커널함수 $K(x_i, x_j)$를 사용하여 최종적으로 아래의 문제로 간략화 된다.

$$\max -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)K(x_i, x_j)+\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)y_i \tag{6}$$
$$-\sum_{i=1}^{l}(\alpha_i+\alpha_i^*)\epsilon$$
$$\text{subject to } \sum_{i=1}^{l}(\alpha_i-\alpha_i^*)=0,\ \alpha_i^{(*)}\in[0,1],\ i=1,\dots,m$$

QP(quadratic programming) 형태로 표현된 쌍대문제(6)을 풀면 최적의 α_i, α_i^*가 각각 구해지고, $\alpha_i^{(*)}\neq 0$인 경우만 서포트 벡터가 되며 서포트 벡터에 의한 회귀함수식은 최종적으로 아래와 같이 주어진다. (여기에서, (*)표기는 *가 있는 경우와 없는 경우가 모두 해당됨을 의미한다.)

$$f(x)=\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)K(x_i, x_j)+\beta_0 \tag{7}$$

또한, 최적의 β_0값은 Kaurush-Khun-Tucker의 보상조건에 의해서 얻어진다. 서포트 벡터 회귀는 그림 1과 같은 신경망과 같은 구조를 갖는다.

3. 서포트 벡터 회귀 제어기 설계 및 모의실험

서포트 벡터 회귀 제어기 설계 및 모의실험을 위하여 그림 1과 같이 수레위의 볼을 제어하는 시스템을 고려한다. 이 시스템의 입력은 수레를 움직이는 힘 f이고 출력은 수레의 움

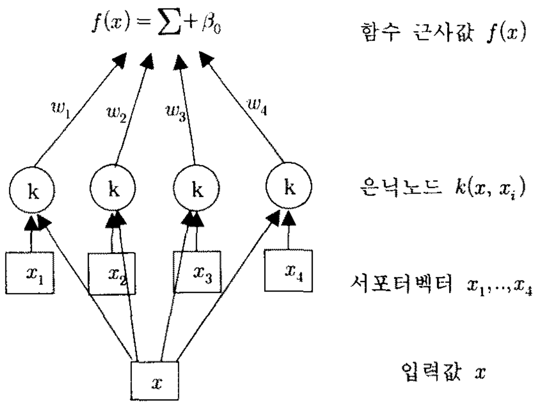


그림 1 서포트 벡터 회귀 구조

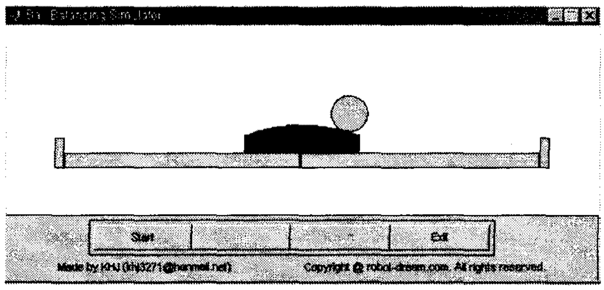


그림 2 수레-볼 제어시스템

직임 $x, \dot{x}$과 수레 위에 있는 볼의 움직임 $\theta, \dot{\theta}$이다. 수레와 볼의 동력학 식을 나누어 구하면 아래와 같다.

$$\text{수레 : } (M+m)\ddot{x}+mlcos\theta\ddot{\theta}-mlsin\theta\dot{\theta}^2+b_x\dot{x}=f \tag{8}$$

$$\text{볼 : } (J+ml^2)\ddot{\theta}+mlcos\theta\ddot{x}-mlgsin\theta+b_\theta\dot{\theta}=0$$

여기에서 M은 수레의 질량(kg), x는 수레의 위치(m), f는 입력 힘(N), J는 볼의 무게 중심에서의 회전관성, m은 진자의 질량(kg), g는 중력 가속도(N/Kg), l은 축에서 볼 무게 중심까지의 길이(m), θ는 진자의 기울어진 각도(rad), b_x는 수레 마찰 계수(N/meter/sec), b_θ는 진자 마찰 계수(N/rad/sec)를 나타낸다. 이러한 복잡한 수식의 해석을 통해 제어기를 설계하는 것은 쉬운일이 아니다.

물리적 특성	특성값
수레 힘	-25 ~ 25N
수레 속력	-5 ~ 5 m/s
수레 위치	-2.5 ~ 2.5 m
볼 속력	-1.25 ~ 1.25 m/s
볼 위치	-50 ~ 50°
수레 곡률	0.5 m
볼 직경	0.055 m
수레 질량	1 kg
볼 질량	1 kg

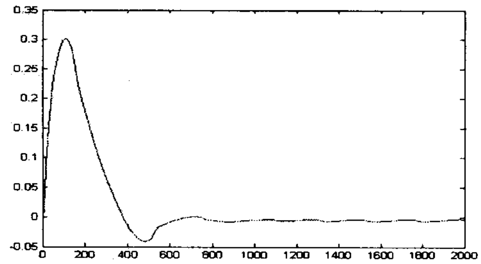
표 1 모의실험을 위한 물리적인 데이터

이러한 단점을 해결하기 위해 지능시스템을 이용한 제어기가 개발되고 있으며 본 논문은 서포트 벡터 회귀 제어기 설계를 위해 실제 실험을 통해서 서포트 벡터 회귀 입력값($x_i \in R^1$)은 수레의 위치, 수레의 속도, 볼의 위치, 볼의 속도와 출력값 ($y_i \in R^1$)은 수레의 움직이는 힘으로 하여 학습데이터를 2000 개를 얻었다. 모의실험에서 적용한 물리적인 데이터를 표 1 에 나타내었다. 모의실험에서 서포트 벡터 회귀 매개변수는 표 2와 같다.

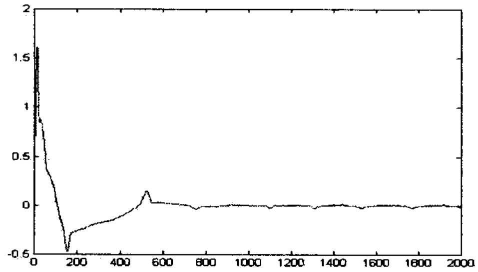
커널	가우시안 RBF 커널
trade-off C	10000
ϵ	0.38
σ^2	1
서포트 벡터 개수	148

표 2 서포트 벡터 매개변수 및 서포터 벡터 개수

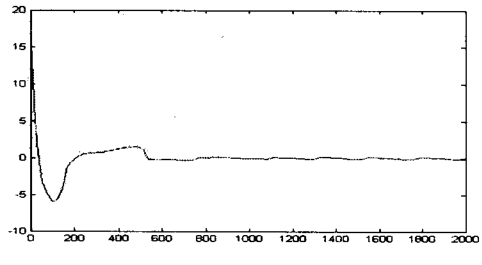
초기값은 수레위치 0, 수레속도 0, 볼 위치 15°, 볼 속력 0이며 실제 모의실험 결과를 그림 3에 나타내었다. 결과에서와 같이 서포트 벡터 회귀 제어기는 모든 상태 변수들이 0으로 수렴하여 안정된 상태를 유지한다.



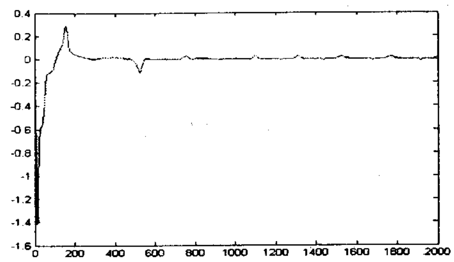
가) 수레 위치(x)의 변화



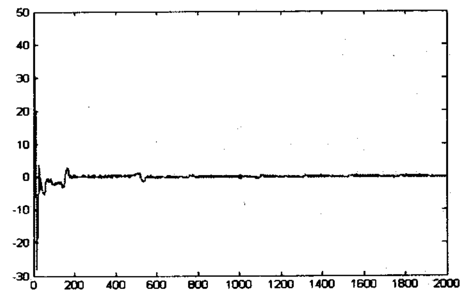
나) 수레 속도($\dot{x}$)의 변화



다) 볼 위치(θ) 변화



라)볼 속도($\dot{\theta}$) 변화



마) 수레에 작용하는 힘의 변화

그림 3 모의실험 결과

5. 결 론

본 논문에서는 1개의 은닉층을 가지는 신경망보다 우수한 성능을 보이는 서포트 벡터 회귀로 시스템의 제어기로 사용 가능성을 확인하였으며 모의실험 결과 성능이 우수함을 보였다.

참 고 문 헌

[1] B. Schölkopf and A.J. Smola : Learning with kernels, , MIT Press, 2002.

[2] C. Burges, "A tutorial on support vector machines for pattern recognition," Knowledge Discovery and Data Mining, vol. 2, no. 2, pp 121-167, 1998.

[3] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," Neuro Colt Technical Report NC-TR-1998-030, Royal Holloway College, University of London, UK, 1998.

[4] B. Schölkopf , J. C. Platt, J. Shawe-Taylor, A. J. Smola and R. C. Williamson, "Estimating the support of a high-dimensional distribution," Neural Computation, vol.13, pp.1443-1471, 2001.