

방사형 기저 함수 기반 다항식 뉴럴네트워크 설계

Design of RBF-based Polynomial Neural Network

김기상*, 진용하**, 오성권***, 김현기****
Ki-Sang Kim, Yong-Ha Jin, Sung-Kwun Oh, Hyun-Ki Kim

Abstract - 본 연구에서는 복잡한 비선형 모델링 방법인 RBF 뉴럴 네트워크(Radial Basis Function Neural Network)와 PNN(Polynomial Neural Network)을 접목한 새로운 형태의 Radial Basis Function Polynomial Neural Network(RPNN)를 제안한다. RBF 뉴럴 네트워크는 빠른 학습 시간, 일반화 그리고 단순화의 특징으로 비선형 시스템 모델링 등에 적용되고 있으며, PNN은 생성된 노드들 중에서 우수한 결과값을 가진 노드들을 선택함으로써 모델의 근사화 및 일반화에 탁월한 효과를 가진 비선형 모델링 방법이다. 제안된 RPNN모델의 기본적인 구조는 PNN의 형태를 이루고 있으며, 각각의 노드는 RBF 뉴럴 네트워크로 구성하였다. 사용된 RBF 뉴럴 네트워크에서의 커널 함수로는 FCM 클러스터링을 사용하였으며, 각 노드의 후반부는 다항식 구조로 표현하였다. 또한 각 노드의 후반부 파라미터들은 최소자승법을 이용하여 최적화 하였다. 제안한 모델의 적용 및 유용성을 비교 평가하기 위하여 비선형 데이터를 이용하여 그 우수성을 보인다.

Key Words : PNN(Polynomial Neural Network), RBFNN(Radial Basis Function Neural Network), FCM(Fuzzy C-Means) Clustering Method, nonlinear multi-variable process data

1장. 서론

자연현상을 분석하고 원하는 정보를 이용하는 것은 물론 생산자동화와 컴퓨터에 의한 생산체제에서의 적절한 작업관리를 위하여 안정된 시스템 모델링은 필요가 아닌 필수인 상황이 도래했다. 그러나 대부분의 현상들은 비선형성을 바탕으로 하고 있으므로 과거의 방법들을 이용해서 상황을 해석하기에는 많은 어려움에 부딪히게 된다. 따라서 시스템의 복잡성과 결과의 불확실성에 대한 분석을 위해 많은 연구가 지속되고 있으며 인간의 언어체계를 표현하는데 효과적인 지능형 모델이 출현하면서, 점진적으로 개선하는 추세이다. 본 연구에서는 입력변수가 증가하면 시스템이 복잡해지고 함수식으로 표현하는데 한계를 드러내는 모델의 설계를 위하여 PNN(Polynomial neural network)[1][2]을 바탕으로 RBF(Radial Basis Function) 뉴럴 네트워크를 적용시킨 RPNN 모델을 제안한다. PNN은 GMDH(Group Method of Data Handling)[3]의 확장형으로 2변수 2차식의 한계를 극복한 모델로 네트워크의 규모 연결성, 뉴런의 표현 형태 등을 자동적으로 결정하는 진화론적 설계구조 모델이다. 비선형성 시스템에서 다항식을 반복적으로 중첩시켜 구축하는데 보다 정확성을 높여 강한 구조를 설계할 수 있다. RBF 뉴럴 네트워크는 하나의 은닉층(hidden layer)만을 가지는 네트워크의 구조로 빠른 학습시간, 일반화, 단순화특징을 가지고 있다. 또한 데이터의 특성을 충분히 반영하기 위하여 Fuzzy

c-means(FCM)클러스터링[4]을 이용하여 구조 동정하여 적용시켰다. 본 논문에서는 RPNN의 구조와 알고리즘을 설명하고, FCM에 의하여 분류되어진 데이터를 사용하여 각 알고리즘의 장점을 최대한 활용한 최적모델을 연구하였다.

2장. 본론

2.1절 다항식 뉴럴네트워크(PNN)의 구조

PNN알고리즘[1][2]은 각 노드의 변수 식들에 여러 가지 형태의 다항식과 입력 변수들의 선택으로 최적의 모델 구조를 얻는다. PNN에서 최적의 모델을 얻기 위해, 회귀 다항식 표 1을 사용한 부분표현식을 구성한다. 다음은 선형회귀분석으로 부분표현식의 계수와 중간변수를 구한다. 중간변수에 의한 입력력 데이터를 구성하여, 다음 계층에서도 이와 같은 형태의 연산을 되풀이하며, 알고리즘의 종료 판정식을 만족할 때까지 계속 반복된다. 알고리즘의 종료 후 얻어지는 부분표현식은 계층적으로 조합하고, 최적의 모델을 얻는다.

표 1. 회귀 다항식의 구조

Type	Polynomial equation
Linear	$y = a_0 + a_1x_i + a_2x_j$
Quadratic	$y = a_0 + a_1x_i + a_2x_j + a_3x_ix_j + a_4x_i^2 + a_5x_j^2$
Modified quadratic	$y = a_0 + a_1x_i + a_2x_j + a_3x_ix_j$

저자 소개

- * 金基祥 : 水原大學校 電氣工學科 碩士課程
- ** 陳庸夏 : 水原大學校 電氣工學科 碩士課程
- *** 吳聖權 : 水原大學校 電氣工學科 正教授·工博
- **** 金賢起 : 水原大學校 電氣工學科 正教授·工博

표 2. 제안된 모델의 성능지수

Polynomial equation	Linear		Quadratic		Modified quadratic	
	PI	EPI	PI	EPI	PI	EPI
AVG±STD	19.92	42.86	27.68	49.50	24.43	40.55
	±2.72	±8.96	±4.07	±14.64	±2.08	±17.41

2.2절 Fuzzy C-means(FCM) Clustering Method

Fuzzy C-means 클러스터링 방법[4][5]은 패턴인식 시스템에 퍼지 집합 이론이 데이터 획득, 특징 추출, 패턴분류 등으로 구성되는 일반적인 인식 시스템에 존재하는 여러 불확실성을 효과적으로 처리할 수 있는 장점 때문에 많이 사용된다. 또한 주어진 출력 데이터의 특징을 해석하여 파악된 출력 데이터의 정보를 기반으로 새로이 입력 데이터를 분할함으로써 기존의 데이터 전처리 방법들보다 더 효율적이고 유기적으로 주어진 입력 데이터의 특성을 파악하여 모델의 구축에 적용할 수 있다. FCM 클러스터링은 그 분류 결과가 소속정도로 표현되므로 유연한 패턴 분류 시스템을 구성할 수 있다는 장점이 있다. FCM 클러스터링은 n개의 벡터 $x_i, i=1, \dots, n$ 의 집합을 c개의 퍼지 그룹들로 분할하고, 비 유사성 측정의 비용함수가 최소가 되는 것과 같은 각각의 그룹 안에서 클러스터의 중심을 찾는다. FCM 알고리즘은 유사척도로서 특정 내적 노름자(Inner product norm metric)를 사용하여 정의된 목적 함수가 근사적 최솟값이 되도록 퍼지분할 U와 클러스터의 중심값 V를 반복 계산법에 의해 구하는 최적화 퍼지 클러스터링 알고리즘으로써 목적 함수는 식 1과 같다.

$$Min z(U, V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m (d_{ik})^2, \quad d_{ik}^2 = \| \vec{x}_k - \vec{v}_i \|^2 \quad (1)$$

- 여기서, x : 임의의 유한 데이터 집합
- V : 클러스터의 중심벡터
- U : 퍼지 분할
- m : 하중 파라미터
- c : 클러스터의 수
- n : 데이터의 수
- $\| \cdot \|$: 내적 노름자

2.3절 퍼지 c-means를 이용한 RPNN의 설계

본 논문에서 제안한 다항식 RPNN 모델은 3층의 PNN 모델을 기초로 한다. PNN의 각 노드에 RBF 뉴럴 네트워크[6]를 사용한다. 사용된 RBF 뉴럴 네트워크에서의 커널 함수로는 FCM 클러스터링을 사용하였으며, 각 노드의 후반부 다항식은 표 1과 같은 형태를 가진 다항식 구조로 표현하였다. 그림 1은 RPNN의 일반적인 구조를 도식적으로 나타내었다.

네트워크에서 표현된 각 노드의 개수 k는 식 2와 같이 표현된다.

$$k = \frac{n!}{(n-r)!r!} \quad (2)$$

- 여기서, n : 총 입력변수의 수
- r : 한 노드의 입력변수의 수

각 노드의 입력 및 클러스터의 개수는 2~5개로 설정하여 각각의 노드를 최소자승법에 의해 파라미터를 구하고, 여기서 추정된 계수 a_0, a_1, \dots, a_n 을 각각의 부분표현식에 대입한다. 평가용 데이터를 사용하여 성능지수를 구하고, 하중값을 가진 목적함수로 최적의 입력변수들의 조합을 선정한다. 이 성능지수는 실제 출력 데이터와 모델 출력 데이터의 RMSE

(Root Mean Square Error)로 정밀도를 평가하였다. 성능지수를 평가하여 다음 층으로 넘겨줌으로써, 출력이 곧 입력의 형태로 바뀌어 적용된다. 설정된 종료 판정, 즉 3층까지 모두 수행한 후 가장 좋은 성능지수를 가지는 노드가 최종 출력이 된다.

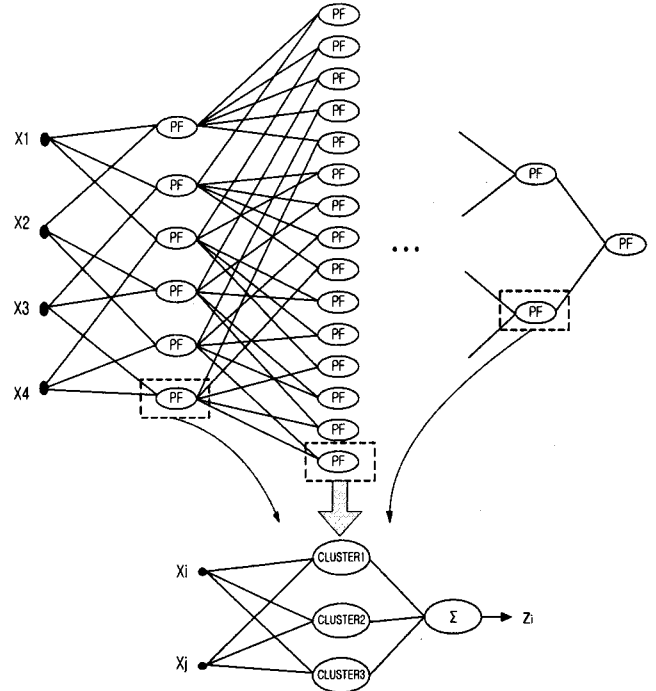


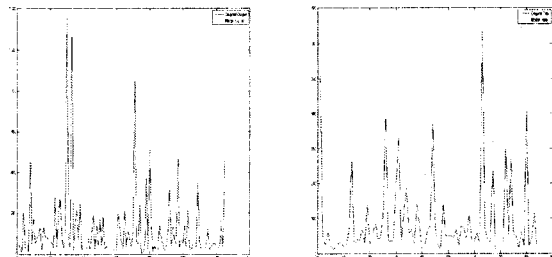
그림 1. 기본 RPNN의 네트워크 구조

3장. 시뮬레이션 및 결과고찰

제안된 네트워크의 구조 그리고 이에 대한 성능의 결과에 대한 우수성을 Computer Hardware Data(CHD)[7]를 이용하여 알아보고자 한다.

CHD는 데이터를 컴퓨터 하드웨어 성능에 따라 상대적인 중앙 처리 장치(CPU)의 성능을 나타내는 데이터이다.[5] 6개의 입력과 상대적인 성능(FRP)을 출력으로 하였다. 209개 데이터를 6 : 4의 비율로 학습 데이터 125개 테스트 데이터 84개로 실험하였다.

다음 표 2는 표 1의 3가지 회귀 다항식 구조를 각각 입력 변수를 2~5, 클러스터 개수를 2~5로 10-fold cross-validation을 이용하여 가장 좋은 데이터별로 나타내었다. 표 2를 보면 회귀 다항식 구조 중 선형 구조가 가장 좋은 성능을 보임을 알 수 있었다. 다음 표 3는 가장 성능이 좋게 나왔던 선형 구조 다항식을 표 2에서와 같은 방법을 이용하여 나타내었다. 다음 그림 2~4는 성능이 가장 좋은 입력변수를 2개 사용하고, 클러스터 개수를 5개 인 모델을 나타낸 것이다.



(b) PI (c) E_PI

그림 2. RPNN의 출력 데이터와 모델 출력 데이터 비교

표 3. 제안된 모델의 성능지수

Number of clusters	Number of inputs							
	2		3		4		5	
	PI	EPI	PI	EPI	PI	EPI	PI	EPI
2	36.14 ±3.12	39.60 ±5.00	32.27 ±5.11	38.13 ±6.34	30.80 ±3.62	42.28 ±13.20	30.40 ±3.73	535.95 ±1.18+e3
3	30.17 ±2.61	41.42 ±7.93	28.91 ±3.87	37.06 ±4.69	25.30 ±3.91	46.26 ±14.79	23.02 ±2.09	1.18+e5 ±2.45+e5
4	27.92 ±2.71	40.32 ±7.02	23.14 ±3.03	41.10 ±8.55	21.10 ±1.85	98.83 ±103.48	18.77 ±2.45	2.66+e6 ±5.53+e6
5	25.63 ±2.78	40.28 ±7.51	19.92 ±2.72	42.86 ±8.96	17.62 ±2.05	214.50 ±467.30	16.16 ±1.98	3.42+e7 ±7.88+e7

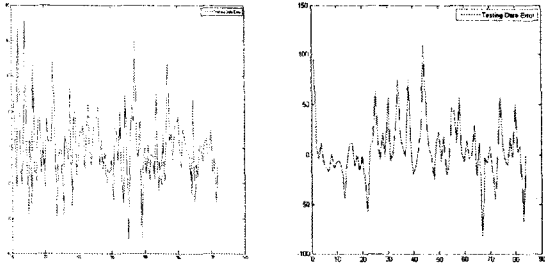


그림 3. RPNN의 출력 데이터와 모델의 출력의 오차

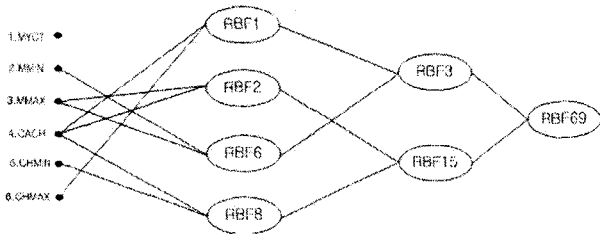


그림 4. 최적 RPNN의 네트워크 구조

4장. 결론

본 연구에서는 비선형 시스템의 수학적 모델을 얻기 위해 RBF를 기반으로 하여 FCM알고리즘을 적용한 다항식 뉴럴 네트워크 구조를 설계하였다. 기존의 PNN 알고리즘은 층을 구성하는 노드의 입력변수의 수와 다항식 차수 그리고 고정된 입력변수 선택을 설계자의 경험에 의해 결정함으로써 객관적이고 주어진 데이터의 특성에 맞는 모델을 제시하지 못하였다. 그러나 본 논문의 RPNN모델은 데이터의 특성을 최대한 반영하기 위하여 클러스터의 수를 다양하게 설정하였고 입력 개수 역시 여러 형태로 적용하였다. 또한 층마다 노드의 선택과 제거를 통하여 최량의 성능지수를 얻을 수 있도록 설계하여 최적의 모델을 얻을 수 있었고, 비선형 공정에 대해 매우 정밀한 예측 성능을 가지고 있음을 보였다.

Computer Hardware Data를 사용함으로써 비선형의 제한된 데이터 수 일지라도 데이터의 특성을 최대한 반영한 구조 및 파라미터 동정에 의한 최적화된 다항식 뉴럴 네트워크를 설계할 수 있음을 알 수 있었다.

감사의 글

본 연구는 경기도의 경기도지역협력연구센터사업의 일환으로 수행하였음 [GGA0801-45700, U-city 보안감시 기술협력센터]. 그리고 본 논문은 2008년도 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (KRF-2008-314-D00376).

참고 문헌

- [1] Oh, S. and Pedrycz, W., 2002, "The Design of Self-Organizing Polynomial Neural Networks", Information Sciences.
- [2] S. K. Oh and Pedrycz, W and B. J. Park, "Polynomial Neural Networks Architecture : Analysis and Design", Computers and Electrical Engineering, 2002
- [3] A. G. Ivahnenko, "Polynomial theory of complex systems", IEEE Trans. On Systems, Man and Cybernetics, Vol. SMC-1, pp. 364-378, 1971.
- [4] W. Pedrycz, "Conditional fuzzy C-means", Pattern Recognition Letters. Vol.17 No.6, 1996. pp.625-632
- [5] 이상경, 장동식, "Genetic 알고리즘과 퍼지 c-Means을 이용한 역전파신경망의 패턴인식성능 개선", 대한군사운영분석학회지, 제 21권, 제 1호, 1995. 6. 30.
- [6] 박호성, 오성권, "Granular Computing에 의한 RBFNN의 구조동정", 한국인터넷정보학회 2008 정기총회 및 추계학술발표대회 제9권 제2호, 2008. 11
- [7] UCI Repository : <http://archive.ics.uci.edu/ml/datasets/Computer+Hardware>