

정보기술응용연구
제1권 제3·4호
1999년 12월

Dempster's Rule of Combination을 이용한 인공신경망간의 결합에 의한 ARMA 모형화

오 상 봉*

요 약

.....

본 논문은 시계열자료의 ARMA 모형화를 위해 계층적(Hierarchical) 문제해결 방식인 인공신경망 기초 의사결정트리분류기상의 인공신경망 구조를 개선하여 지역문제(Local Problem)를 해결하는 복수개의 인공신경망 결과를 Dempster's rule of combination을 이용하여 종합하는 병행적인(Parallel) ARMA 모형화를 위한 방법론을 제시함으로써 의사결정트리분류기에 근거한 방법론의 단점을 보완하였다. 본 논문에서 제시한 ARMA 모형화를 위한 방법론은 세 단계로 구성되어 있다: 1) ESACF 특성 벡터 추출단계; 2) 개별 인공신경망에 의한 부분적 모델링 단계; 3) Conflict Resolution 단계. 제시한 방법론을 검증하기 위해 모의실험용 자료와 실제 시계열자료를 이용하여 제시된 방법론을 검증하였으며 실험결과 기존 연구에 비해 ARMA 모형화의 정확도가 높은 것으로 나타났다.

.....

* 대전대학교 정보통신공학과

1. 서론 및 연구동향

시계열자료(Time Series Data)는 시각의 경과에 따라 일정한 간격으로 측정된 일련의 자료를 의미하는 것으로 그 동안 여러 학문분야의 주요 연구주제의 하나로 다루어져 왔다. 그 동안의 시계열자료에 대한 연구는 과거 자료의 분석을 통해 미래 자료를 예측하려는 노력으로 귀결된다고 해도 과언은 아닐 것이다. 시계열자료의 분석은 시계열자료의 모형화(Identification), 모형내 파라미터의 추정(Parameter Estimation), 그리고 모형의 적합성 분석(Diagnostic Checking)으로 나누어진다.

특히, Box and Jenkins[4]가 제시한 ARMA(Autoregressive Moving Average) 모형은 학계나 실무에서 가장 널리 이용되는 시계열자료에 대한 분석 툴로써 통계분야에서 오랜 전통을 갖는 연구분야의 하나이며 많은 통계학적인 시계열 모형화 방법이 제시되었다([2], [3], [6], [7], [13], [14], [17], [18], [19]). 그러나 이러한 통계적 접근방법은 이론적 측면에서의 우수성에도 불구하고 복잡한 통계치와 시계열자료에 내포된 잡음으로 인해 실용적인 측면에서 제한성을 가질 수밖에 없었다. 복잡한 통계적 이론은 시계열 모형화를 하고자 하는 사용자들에게 어려움을 주고 있으며 동시에 시계열자료에 내포된 잡음은 이론적 통계치와 실제 시계열자료로부터 도출된 통계치 사이에 차이를 낳게 되어 사용자들은 시계열 모형화를 위해 실제 통계치와 이론적 통계치사이의 유사성에 대한 판단을 해야만 한다는 것이다.

위에서 지적한 통계적 방법론의 문제점을 보완하기 위해 시계열 모형화를 위한 많은 인공 지능적인 방법들이 제시되었다. 즉, 이론적 패턴과 잡음으로 인한 실제 패턴사이의 유사성 판단은 패턴인식의 문제이므로 패턴인식에서 상당한 유용성을 보인 인공 지능적인 방법들이 단독으로 혹은 통계적 방법과 결합된 형태로 연구되기 시작하였다. Lee and Park([11], [12])의 연구는 의사결정트리분류기(Decision Tree Classifier)와 패턴인식방법을 결합하여 시계열자료의 ARMA 모형화를 시도하였다. Jhee et al.[8]은 Box and Jenkins[4]에서 제시된 ARMA 모형화 과정에서 AR 차수와 MA 차수를 결정하는데 인공신경망을 이용하여 Box and Jenkins[4] 방법의 불확실성을 줄이고자 하였다.

다음의 연구들은 Tsay and Tiao[17]가 제시한 ESACF(Extended Sample AutoCorrelation Function)을 이용하여 ARMA 모형화를 시도한 연구들이다. Lee

and Jhee[9]는 시계열자료로부터 도출된 ESACF을 2단계에 걸쳐 인공신경망을 적용한 방법론을 제시하고 있다. 첫 번째 단계에서는 ESACF상에 존재하는 잡음을 제거하고, 이렇게 잡음이 제거된 ESACF를 이용하여 ARMA 모형화를 시도하였다. Lee and Oh[10]와 오상봉, 이진창[1]에서는 시계열자료로부터 도출된 ESACF을 인공신경망과 의사결정트리분류기를 결합한 인공신경망 기초 의사결정트리분류기(Neural Network-Driven Decision Tree Classifier: NNDTC)에 적용하여 ARMA 모형화의 정확도를 높이고자 하였다.

한편, 시계열 모형화 방법론과는 직접적인 관련은 없지만 본 논문의 방법론과 관련하여 패턴인식을 위해 복수개의 패턴분류기를 적용한 연구를 들 수 있다([5], [20], [21]). 이들은 특정 문제를 위해 복수개의 패턴분류기를 훈련시킨 후 입력자료를 복수개의 패턴분류기에 적용하고 이들로부터 나온 결과를 특정 연산자를 이용해 결합하는 방법을 제시하고 있다.

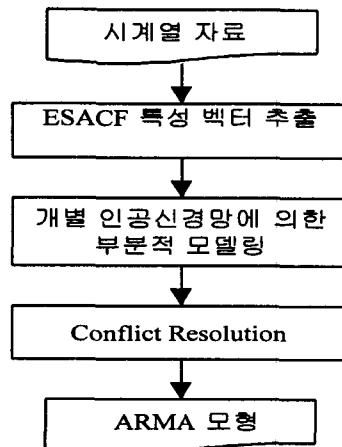
본 논문은 시계열자료의 ARMA 모형화를 위해 계층적(Hierarchical) 문제해결 방식인 인공신경망 기초 의사결정트리분류기상의 인공신경망 구조를 개선하여 개선된 인공신경망을 이용해 병행적인(Parallel) 문제해결 방식의 ARMA 모형화를 위한 방법론을 제시하고자 한다. 또한 모의실험용 자료와 실제 시계열자료를 이용하여 제시된 방법론을 검증할 것이다. 본 논문의 구성은 2장에서 인공신경망 기초 의사결정트리분류기와 비교하여 ARMA 모형화 방법론을 설명할 것이고 3장에서는 제시된 방법론을 이용한 실험결과가 제시될 것이다. 4장에서는 실제 시계열자료를 대상으로 실험한 결과를 보일 것이고 5장에서 결론 및 향후 연구방향에 대해 언급하기로 한다.

2. 방법론

본 논문에서 제시하고자 하는 방법론은 [그림-1]과 같이 세 단계로 구성되어 있다. 첫 번째 단계는 시계열자료로부터 ESACF 특성 벡터를 추출하는 과정이다. 두 번째 단계는 추출된 ESACF를 입력자료로 이용하여 학습된 개별 인공신경망으로부터 ARMA 모형화를 위한 부분적인 모델링을 함으로써 분할된 지역문제에 대한 지역해를 구하는 과정이다. 세 번째 단계는 개별 인공신경망을 통해 얻은 지역해를 종합하여 전체문제에 대한 해를 구하는 과정으로 각각의 개별 인공신경망간에 존재할 수 있는 모순을 해결하여 하나의 ARMA 모형으로 일원화하게 된다.

2.1 ESACF 특성 벡터 추출단계

ESACF는 Tsay and Tiao[17]에 의해 제안된 통계량으로 잡음이 없는 순수한 ARMA(2,3) 모형의 경우, [표-1]과 같은 삼각형 구조의 패턴을 보이게 된다. 즉, ARMA모형의 AR 차수와 MA 차수에 해당되는 점에서 삼각형의 꼭지점이 형성된다. [표-1]에서 1로 표시된 부분은 잔차간의 자기상관계수가 표준편차의 두 배수 안에 존재하는 것을 의미하는 것이고 0은 잔차간의 자기상관계수가 표준편차의 두 배수밖에 존재하는 경우를 의미한다. 그러나 서론에서도 밝힌 바와 같이



[그림-1] ARMA 모형화를 위한 방법론

[표-1] 순수한 ARMA(2,3) 모형의 ESACF 특성 벡터

AR \ MA	0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	1	1	1	1	1	1	1
3	0	0	0	0	1	1	1	1	1	1
4	0	0	0	0	0	1	1	1	1	1
5	0	0	0	0	0	0	1	1	1	1
6	0	0	0	0	0	0	0	1	1	1
7	0	0	0	0	0	0	0	0	1	1
8	0	0	0	0	0	0	0	0	0	1
9	0	0	0	0	0	0	0	0	0	0

시계열자료에 내포된 잡음으로 인해 삼각형의 모양이 왜곡된 형태로(즉 0이 1로, 혹은 1이 0으로 바뀌게 됨) 나타나게 되어 삼각형의 꼭지점이 어느 위치에 존재하는가를 판단하기 어려운 것이 현실이다. 이러한 어려움을 해결하기 위해 많은 연구에서는 패턴분류기를 이용하여 순수한 통계적인 방법의 문제점을 해결하고자 하였고 최근에는 패턴분류기로 인공신경망을 많이 이용하고 있다.

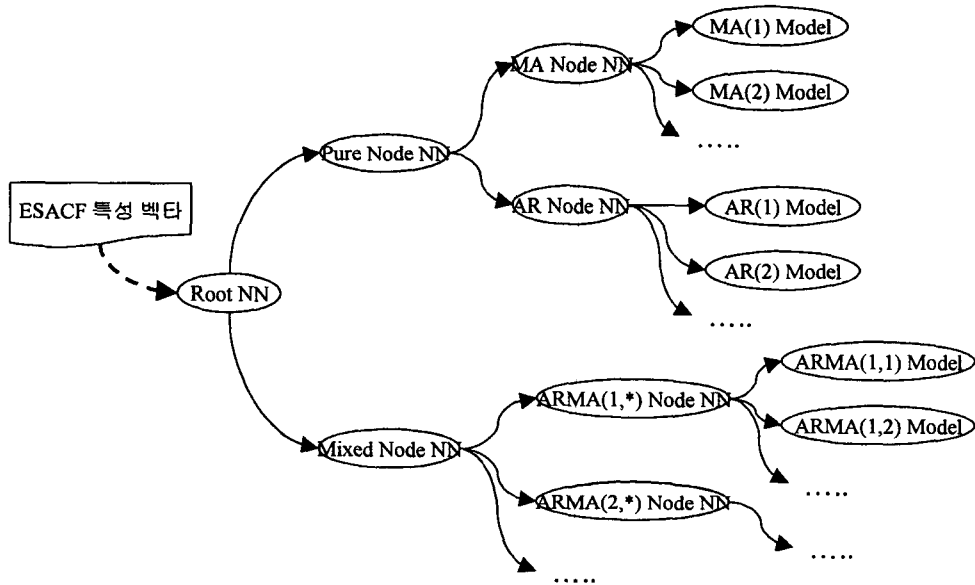
2.2 개별 인공신경망에 의한 부분적 모델링 단계

Lee and Oh[10], 오상봉, 이진창[1]에서 제시한 NNDTC는 [그림-2]에서 알 수 있는 바와 같이 의사결정트리분류기의 노드에 인공신경망을 부착하여 해당 인공신경망의 결과 값에 의해 하위노드가 결정되는 계층적인 의사결정구조를 갖는 ARMA 모형화 방법론으로 요약할 수 있다. 즉 MA Node NN은 MA 모형 중에서 MA 차수를 결정할 수 있는 인공신경망으로 결과 값에 따라 어느 한쪽의 노드로 이동하게 되고 도착된 leaf 노드가 주어진 시계열자료의 ARMA 모형이 되는 것이다(예를 들면 ARMA(0,1) 즉 MA(1) 모형).

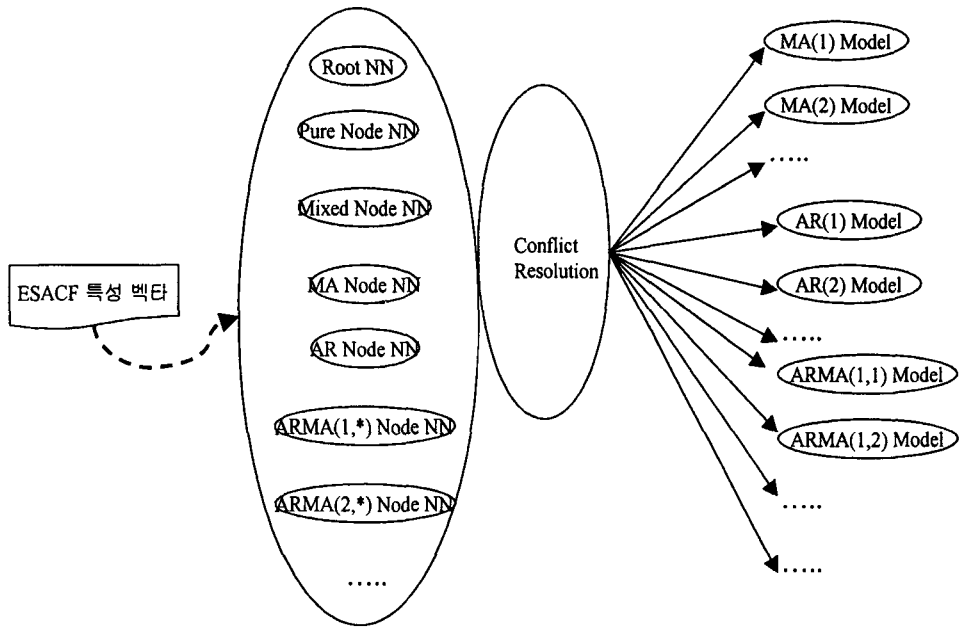
비록 NNDTC가 ESACF 특성 벡터를 이용하여 인공신경망과 의사결정트리분류기의 장점을 결합했다고 할 수 있으나 계층적인 의사결정구조로 인해 하위노드에서는 상위노드에서의 잘못된 분류를 복구할 수 없다는 점에서 아쉬움이 남는다고 할 수 있다(상위노드에 부착된 인공신경망의 가중치가 하위노드의 인공신경

망의 가중치 보다 우선하는 Lexicographic Ordering 구조). 본 논문에서는 NNDTC의 이와 같은 문제점을 해결하고자 의사결정트리분류기상의 인공신경망을 개별 의사결정주체로 인식하여 각각의 인공신경망은 자기에게 할당된 지역문제를 해결하도록 하고 동시에 이들 인공신경망에 같은 가중치를 부여하여 각각의 인공신경망의 지역해를 종합하는 방식으로 계층적인 의사결정구조를 병행구조로 변경하였다([그림-3] 참조).

이를 위해 기존 NNDTC상의 해당 노드에 부착된 인공신경망의 구조를 변경해야 하고 각각의 인공신경망의 결과를 종합할 Conflict Resolution Mechanism이 고안되어야 할 것이다. Conflict Resolution은 다음 절에서 다루기로 하고 본 절에서는 인공신경망의 구조를 어떻게 변경할 지를 설명하기로 한다. NNDTC상의



[그림-2] NNDTC 방법론



[그림-3] ARMA모형화 방법론 상세도

노드에 부착된 인공신경망의 입력노드는 ESACF 특성 벡터에 의해 결정되고 출력노드 수는 해당 노드에 연결된 하위노드의 수로 결정된다. 따라서 계층적인 의사결정트리분류기에서는 상위노드의 잘못을 하위노드에서 복구할 수 없게 되는 것이다. 본 논문에서 제안한 방법론에 이용할 인공신경망의 입력노드의 수는 NNDTC와 마찬가지로 ESACF 특성 벡터에 의해 결정되지만 출력노드의 수는 NNDTC의 해당 인공신경망의 출력노드의 수보다 한개 많은 구조를 가지게 된다. 추가된 1개의 노드는 Other Model을 의미하는 것으로 NNDTC 상의 상위노드에서 잘못 분류된 경우에 이를 복구하기 위해 추가된 노드이다.

MA Node NN을 예로 들어보자. NNDTC상의 인공신경망은 5개의 출력노드를 갖게 된다(ARMA(5,5)까지로 모형을 제한한 경우). 그러나 본 논문에서 제안한 인공신경망은 6개의 출력노드를 갖게 되고 추가된 노드는 MA(1), MA(2), ..., MA(5) 이외의 모형일 경우에 출력값 1을 갖는 노드이다. ARMA(1,*) Node NN의 경우에는 6개의 출력노드를 갖게 되고 추가된 출력노드는 ARMA(1,1)

,...ARMA(1,5) 이외의 모형에 해당되는 ESACF 특성 벡터의 입력에 대해 출력값 1을 갖도록 학습될 것이다. 나머지 노드에 부착된 인공신경망도 NNDTC상의 인공신경망보다 한개 많은 출력노드를 갖게 되고 추가된 출력노드는 NNDTC상의 해당 노드에 연결된 하위노드에 없는 모형을 의미하는 것이다. 다만, Root NN의 경우에는 NNDTC와 마찬가지로 2개의 출력노드를 갖게 되는 데 이는 본 논문에서 모형화 하고자 하는 모든 시계열자료의 ARMA모형이 Pure 모형(AR 혹은 MA모형)이거나 Mixed 모형이기 때문이다.

2.3 Conflict Resolution 단계

개별 인공신경망의 결과를 종합하기 위해 본 논문에서는 Dempster's rule of combination[16]을 이용하고자 한다. 본 논문에서 Conflict Resolution의 방법으로 Dempster's rule of combination을 채택한 이유는 연산의 단순성과 연산의 교환성, 즉 인공신경망의 결과를 결합하는 데 있어 결합에 참여하는 인공신경망의 순서와 무관하게 결과가 같다는 것이다. 물론 Dempster's rule of combination이외에도 여러 가지 방법을 생각할 수 있고 어느 방법을 택하느냐에 따라 인공신경망의 결합결과는 달라질 수 있을 것이다. 어느 연산방법을 이용하는 것이 효과적이나에 대한 연구는 향후의 과제로 남겨 두기로 한다.

패턴인식을 위해 복수개의 패턴인식기를 적용한 연구([5],[20],[21])와 본 논문의 방법론을 비교해 보면 복수개의 패턴인식기의 결과를 결합한다는 점에서 유사성이 있으나, 이들 연구에서는 입력자료의 표현양식을 변화시켜 얻은 각각의 자료에 대해 패턴인식기를 적용하여 얻은 인식결과를 결합하는 방식을 취하고 있다. 즉, 복수개의 패턴인식기로 같은 문제에 대해 여러 번 해를 구하고 이들의 결과를 결합했다는 것이다. 본 논문에서는 전체 문제를 여러 개의 지역문제로 분할하여 각각의 지역문제에 대해 패턴인식기를 적용하였고 이들 결과를 결합하여 전체 문제에 대한 해를 구한다는 점에서 이들 연구의 방법론과 차이가 있다고 할 수 있다(같은 문제를 다른 패턴인식기를 이용하여 여러 번 해를 구하고 이들 결과를 결합하느냐 vs. 전체 문제를 분할하여 분할된 문제 각각에 대해 패턴인식기를 적용하고 이들 결과를 결합하여 전체 문제에 대한 해를 구하느냐).

Conflict Resolution 단계에서 개별 인공신경망에 부여된 가중치는 모두 같고, 또한 병행적 의사결정구조로 인해 일부 인공신경망의 잘못된 분류는 해당 인공신경망이 NNDTC상의 상위노드에 부착된 인공신경망에 대응되는 인공신경망이라

하더라도 다른 인공신경망의 결과에 의해 복구되거나 최소화될 수 있다. 가령 Root 노드에서 잘못 분류된 경우, NNDTC에서는 이를 복구할 수 없으나(Mixed 모형이 Root 노드에서 Pure 모형으로 분류되면 하위노드에서는 이를 복구하지 못하고 결국 AR 모형이나 MA모형으로 분류됨) 본 논문의 방법론에서는 하위노드의 인공신경망에 대응되는 인공신경망들의 결과에 따라 상위노드에 대응되는 인공신경망의 잘못이 복구되거나 최소화될 수 있다는 것이다(Root 노드에 대응되는 인공신경망이 Mixed 모형을 Pure 모형으로 잘못 분류할지라도 하위노드에 대응되는 인공신경망이 이를 강력하게 부인하는 경우 Pure 모형으로 복구되거나 Pure 모형에 가까운 모형으로 모형화될 수 있음).

3. 실험결과

3.1 인공신경망 구조

대부분의 시계열자료는 AR 차수와 MA 차수가 동시에 5를 넘지 않는 (ARMA(5,5) 모형 안의 모형) 것이 일반적이므로, 본 논문에서 ARMA 모형화를 위해 총 10개의 인공신경망이 필요하며 이들의 구조는 [표-2]에 요약되어 있다. 본 논문에서는 1개의 은닉노드층을 갖는 3 계층구조를 갖는 인공신경망을 이용하였으며 은닉노드 수는 Lee and Oh[10] 와 오상봉,이건창[1]의 연구에서와 같이 인공신경망의 성과가 특별히 나쁘지 않는 한 입력노드 수와 같게 함으로써 인공신경망의 구조를 가능한 단순하게 하였다. 인공신경망의 입력자료에 해당되는 ESACF 특성 패턴을 10 X 10 행렬을 이용하였다(입력노드 수는 100개). 출력노드 수에 대한 자세한 내용은 2.2절에 설명되어 있다. 또한, 인공신경망의 학습을 위해 Bipolar Sigmoid 함수를 이용한 역전파 학습 알고리즘을 채택하였다[15].

3.2 데이터

인공신경망을 학습시키기 위해 네 종류의 학습 데이터를 준비하였다: 원형 데이터, 10% 잡음 데이터, 20% 잡음 데이터, 30% 잡음 데이터. 원형 데이터는 잡음이 섞이지 않은 ESACF 특성 패턴을 의미하고 잡음 데이터는 원형 데이터의

[표-2] 인공신경망 구조

인공신경망	입력노드 수	은닉노드 수	출력노드 수
Root NN	100	100	2
Pure Node NN	100	100	3
Mixed Node NN	100	100	6
AR Node NN	100	100	6
MA Node NN	100	100	6
ARMA(1,*) Node NN	100	100	6
ARMA(2,*) Node NN	100	100	6
ARMA(3,*) Node NN	100	100	6
ARMA(4,*) Node NN	100	100	6
ARMA(5,*) Node NN	100	100	6

100개의 이진수중에서 해당 비율만큼 무작위로 선택하여 선택된 이진수에 대해 이진변환을 한 데이터이다. 다만, ESACF 특성 패턴 안의 삼각형 구조가 심하게 왜곡되는 것을 막기 위해 Lee and Jhee [9], Lee and Oh[10], 오상봉,이건창[1]의 연구와 같이 삼각형안의 잡음이 ARMA(p,q)모형의 경우, $12-p-q$ 를 넘지 않도록 하였다. 데이터 Set은 35개 ARMA모형 각각에 대해 해당 잡음비율을 갖는 한개의 ESACF 특성패턴으로 구성되어 총 35개의 데이터를 포함하고 있으며 각각의 잡음비율에 대해 3개씩의 학습데이터 Set을 준비하였다.

학습된 인공신경망의 성과를 평가하기 위해 두 종류의 테스트 데이터를 준비하였다. 하나는 개별 인공신경망이 제대로 학습되었는 지를 평가(Local 테스트)하기 위한 데이터이고 다른 하나는 10개의 인공신경망으로부터 나온 결과를 종합한 결과 주어진 ESACF 특성 패턴을 올바르게 ARMA 모형화했는 지를 평가(Global 테스트)하기 위한 데이터이다. 두 가지 모두 각각의 잡음수준에 대한 10개의 테스트 데이터 Set으로 구성되었으며 각각의 테스트 데이터 Set은 35개 모형 각각에 대해 해당 잡음비율을 갖는 한개의 ESACF 특성 패턴으로 구성되어 총 35개의 데이터로 구성되었다.

3.3 학습 및 개별 인공신경망 테스트(Local Test)

인공신경망의 학습을 위해 학습율(Learning rate)과 Momentum을 각각 0.25와 0.9로 하였으며 MSE(Mean Squared Error)가 0.005 미만인 경우에 학습이 종료 되도록 하였다. 인공신경망의 학습전략으로는 반복학습전략([1],[10])을 채택하였다. 즉 원형 데이터 Set으로 학습된 인공신경망이 10%, 20%, 30% 잡음비율을 갖는 테스트 데이터 Set에 대해 평균 95% 이상의 정확도를 보이면 학습을 중단하고 그렇지 않은 경우, 원형 데이터 Set과 3개의 10% 잡음비율의 학습데이터 Set을 이용하여 재학습시키고 다시 테스트를 하게 된다. 이러한 과정을 30% 잡음비율의 학습데이터 Set까지 반복적으로 적용하는 것이다(자세한 내용은 [1],[10] 참조). 각각의 인공신경망에 반복학습전략을 적용한 결과가 [표-3]에 요약되어 있다([표-3]의 학습단계의 데이터 Set의 +는 이전의 데이터에 합하여 학습을 했다는 의미이다).

3.4 ARMA 모형화 정확도 테스트(Global Test)

본 논문에서 제안한 방법론을 이용하여 ARMA 모형화를 시도한 결과 10%, 20%, 30% 잡음수준의 테스트 데이터에 대해 각각 97.7%, 92.9%, 89.1%의 정확도를 보이고 있어 기존의 다른 연구에서 보인 결과와 비교할 때 상당히 우수한 것으로 나타났다. [표-4]에 나타난 숫자는 전체 데이터(35개)중에서 정확히 모형화한 데이터의 수를 나타내고 있다. 물론 같은 데이터를 이용하여 각각의 연구 결과를 비교하는 것이 더욱 정확하겠지만 각각의 연구에서 사용한 잡음 데이터들이 원형 ESACF 특성 패턴 행렬 상에서 그 위치를 무작위로 선정하여 해당 위치의 값을 이전 변환하였다는 것을 생각해 볼 때 비록 같은 데이터는 아닐 지라도 그 의미는 크다고 할 수 있다([1]의 연구와는 같은 데이터로 비교하였음).

4. 실제 데이터를 이용한 ARMA 모형화

본 장에서는 시계열자료의 ARMA 모형화를 위해 많이 이용되는 실제 데이터에 본 논문에서 제안한 방법론을 적용하여 제시된 방법론의 정확성을 검토하고자 한다. 먼저 본 논문의 방법론을 Caffeine Data([17])에 적용하여 ARMA 모형화 과정을 설명하기로 한다.

[표-3] 개별 인공신경망의 학습 및 Local 테스트 결과

인공신경망	학습단계			테스트 단계		
	데이터 Set	# of epochs	MSE	10%잡음 데이터	20%잡음 데이터	30%잡음 데이터
Root NN	원형	25	0.0007	0.96	0.8457	0.7971
	+10% 잡음	10	0.0012	0.9628	0.8628	0.8781
	+20% 잡음	31	0.0006	0.9743	0.9086	0.9124
	+30% 잡음	21	0.0021	0.9743	0.9343	0.9286
Pure Node NN	원형	23	0.0049	0.9486	0.8914	0.6943
	+10% 잡음	11	0.0046	0.9657	0.90	0.7057
	+20% 잡음	20	0.0044	0.9543	0.92	0.8086
	+30% 잡음	42	0.0044	0.94	0.9257	0.9257
Mixed Node NN	원형	77	0.0011	0.9743	0.8743	0.7171
	+10% 잡음	6	0.0009	0.9714	0.94	0.7943
	+20% 잡음	7	0.0034	0.96	0.9457	0.8143
	+30% 잡음	27	0.0042	0.9543	0.9286	0.9057
AR Node NN	원형	18	0.0034	0.9829	0.9714	0.8314
	+10% 잡음	6	0.0027	0.9743	0.9543	0.8771
	+20% 잡음	12	0.0049	0.9686	0.9714	0.9257
MA Node NN	원형	31	0.0043	0.9971	0.9943	0.9514
ARMA(1,*) Node NN	원형	49	0.0041	0.9857	0.9229	0.90
	+10% 잡음	10	0.0039	0.9857	0.9543	0.9086
ARMA(2,*) Node NN	원형	57	0.0044	0.9943	0.9257	0.8686
	+10% 잡음	11	0.0024	0.9571	0.9314	0.88
	+20% 잡음	13	0.0041	0.9886	0.9371	0.92
ARMA(3,*) Node NN	원형	54	0.0045	0.9743	0.9229	0.9057
	+10% 잡음	12	0.0027	0.9714	0.9429	0.9343
ARMA(4,*) Node NN	원형	77	0.0024	0.9771	0.94	0.9257
ARMA(5,*) Node NN	원형	63	0.0020	0.9943	0.96	0.9143

[표-4] ARMA 모형화 정확도 테스트 결과

테스트 데이터 Set ID	잡음비율		
	10%	20%	30%
1	35/35	32/35	32/35
2	33/35	33/35	30/35
3	34/35	32/35	30/35
4	34/35	33/35	33/35
5	34/35	31/35	32/35
6	35/35	31/35	31/35
7	35/35	34/35	32/35
8	33/35	33/35	30/35
9	34/35	32/35	33/35
10	35/35	34/35	29/35
평균	97.7%	92.9%	89.1%
Lee and Jhee[9]	96.7%	89.7%	80.8%
Lee and Oh[10]	95.4%	90.9%	86.9%
오상봉, 이진창[1]	96.6%	90.9%	86.6%

4.1 ESACF 특성 패턴 추출

Caffeine Data에 대한 ESACF 특성패턴은 [표-5]에 나타나 있다. [표-5]에서 알 수 있는 바와 같이 ESACF 특성 패턴 행렬에 존재하는 잡음으로 인해 Caffeine Data에 대한 ARMA 모형을 찾아내기는 쉽지 않은 일이다. Caffeine Data에 대한 실제 ARMA 모형은 ARMA(2,5)로 알려져 있다.

4.2 인공신경망의 출력 도출

Caffeine Data에 대한 10개의 인공신경망의 출력노드의 의미와 출력값은 [표-6]에 나타나 있다. Pure Node NN에서의 Other Model의 의미는 AR Model도 아니고 MA Model도 아닌 즉 Mixed Model의 의미이고 ARMA(1,*) Node NN의 Other의 의미는 ARMA(1,i), i=1,2,3,4,5 모형이 아닌 Pure Model (AR(i), MA(i) i=1,2,3,4,5) 이거나 ARMA(i,j), i=2,3,4,5; j=1,2,3,4,5 중 어느 하나일 가능성을 의미하는 것이다.

4.3 Conflict Resolution 단계

10개의 인공신경망의 출력값을 얻게 되면 각각의 인공신경망의 추론결과를 종합하여 가장 가능성이 높은 ARMA 모형을 제안하는 과정이 뒤따르게 된다. 본 논문에서는 앞서도 언급한 바와 같이 Dempster's Rule of Combination을 이용하고자 한다. 10개의 인공신경망의 결과를 종합하는 과정을 모두 소개할 수는 없고 Root NN 과 Pure Node NN의 결합을 예로 제시하고 나머지는 같은 과정의 반복이므로 최종 결과만을 제시하기로 한다.

[표-5] Caffeine Data에 대한 ESACF 특성 패턴

AR \ MA	0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0	0
1	0	1	1	1	0	1	0	0	1	1
2	1	1	1	1	0	1	1	1	1	1
3	0	0	1	1	0	0	1	1	1	1
4	0	0	1	1	0	0	0	1	1	1
5	1	0	1	0	0	0	1	1	0	1
6	0	0	0	0	0	1	1	1	0	1
7	0	1	1	0	0	1	1	1	0	0
8	1	1	1	0	0	1	1	1	1	1
9	1	1	1	0	0	1	1	1	1	0

[표-6] Caffeine Data에 대한 인공신경망 출력노드 및 출력값

인공신경망	출력노드 의미 및 출력값					
	Pure Model			Mixed Model		
Root NN	0.222228			0.777958		
Pure Node NN	AR Model		MA Model		Other Model	
	0.243233		0.003340		0.839799	
Mixed Node NN	ARMA1*	ARMA2*	ARMA3*	ARMA4*	ARMA5*	Other
	0.092096	0.937064	0.000436	0.002424	0.024460	0.000145
AR Node NN	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)	Other
	0.039911	0.186981	0.064692	0.020591	0.062584	0.776179
MA Node NN	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)	Other
	0.012911	0.002739	0.054492	0.036903	0.008895	0.996584
ARMA(1,*) Node NN	ARMA11	ARMA12	ARMA13	ARMA14	ARMA15	Other
	0.069304	0.009979	0.000508	0.322652	0.015382	0.983890
ARMA(2,*) Node NN	ARMA21	ARMA22	ARMA23	ARMA24	ARMA25	Other
	0.528933	0.007138	0.011634	0.032823	0.562910	0.246899
ARMA(3,*) Node NN	ARMA31	ARMA32	ARMA33	ARMA34	ARMA35	Other
	0.000491	0.002305	0.000037	0.001172	0.002270	0.999999
ARMA(4,*) Node NN	ARMA41	ARMA42	ARMA43	ARMA44	ARMA45	Other
	0.016549	0.017832	0.005910	0.001368	0.000983	0.998795
ARMA(5,*) Node NN	ARMA51	ARMA52	ARMA53	ARMA54	ARMA55	Other
	0.022437	0.265258	0.026404	0.048509	0.018313	0.522721

먼저 Root NN으로부터 Caffeine Data에 대해 Pure Model일 가능성(Credibility: CR로 표기)이 0.222228이고 Mixed Model일 가능성이 0.777958임을 알 수 있다. 이 두 개의 값을 Normalization한 결과는 아래와 같다.

$$CR(\text{Pure Model}) = 0.222228 / (0.222228 + 0.777958) = 0.2222$$

$$CR(\text{Mixed Model}) = 0.777958 / (0.222228 + 0.777958) = 0.7778$$

같은 방법으로 Pure Node NN으로부터 아래의 내용을 도출할 수 있다.

$$CR(\text{AR Model}) = 0.2239$$

$$CR(\text{MA Model}) = 0.0031$$

$$CR(\text{Other}=\text{Mixed Model})=0.7730$$

위의 두 결과를 Dempster's Rule of Combination을 이용하여 결합하면 [표-7]과 같다. [표-7]의 결과를 Normalization을 한 결과, 다음을 얻을 수 있다([표-7]안의 \emptyset 는 공집합을 의미함).

$$CR(\text{AR Model})=0.0497/(0.0497+0.0007+0.6013)=0.0763$$

[표-7] Caffeine Data에 대한 Root NN과 Pure Node NN의 결합

	CR(AR Model) (0.2239)	CR(MA Model) (0.0031)	CR(Other Model) (0.7730)
CR(Pure Model) (0.2222)	CR(AR Model) (0.0497)	CR(MA Model) (0.0007)	\emptyset
CR(Mixed Model) (0.7778)	\emptyset	\emptyset	CR(Mixed Model) (0.6013)

$$CR(\text{MA Model})=0.0007/(0.0497+0.0007+0.6013)=0.0010$$

$$CR(\text{Mixed Model})=0.6013/(0.0497+0.0007+0.6013)=0.9226$$

위의 결과와 나머지 8개의 인공신경망의 결과를 결합한 최종 결과 중 가능성이 높은 순서로 ARMA 모형을 제시하면 다음과 같다.

$$\text{ARMA}(2,5)=0.4859$$

$$\text{ARMA}(2,1)=0.4566$$

$$\text{ARMA}(2,4)=0.0283$$

$$\text{ARMA}(2,3)=0.0100$$

따라서 Caffeine Data의 ARMA 모형으로 알려진 ARMA(2,5)모형을 제시하고 있어 본 논문의 방법론이 Caffeine Data를 제대로 ARMA 모형화하고 있다는 것을 알 수 있다. 또 다른 실제 데이터에의 적용 결과는 [표-8]에 요약되어 있다.

5. 결론

본 논문은 시계열자료의 ARMA 모형화를 위한 NNDTC를 개선하여 NNDTC가 가지는 상위노드에서의 잘못된 분류를 하위노드에서 복구할 수 없다는 문제점을 개선하기 위해 계층적 문제해결 방식을 병행적 문제해결 방식으로 개선한 방법론을 제안하였다. 이를 위해 NNDTC상의 인공신경망의 구조를 개선하였으며 복수개의 인공신경망의 결과를 종합하기 위한 Conflict Resolution의 방법으로 Dempster's rule of combination을 이용하였다. 본 논문에서 제시하는 ARMA 모형화 방법론은 1) ESACF 특성 벡터 추출단계, 2) 개별 인공신경망에 의한 부분적 모델링 단계, 3) Conflict Resolution 단계로 구성되어 있다.

또한, 복수개의 패턴인식기를 이용하여 같은 문제에 대한 복수개의 해를 구하고 이들을 결합한 연구와는 달리 본 논문의 방법론은 전체 문제를 복수개의 지역문제(Local Problem)로 분할하고 각각의 지역문제에 대한 인공신경망 결과를 병행적인 방법으로 결합하여 전체문제의 해를 구한다는 점에서 이들의 연구방법론과 차이가 있다고 할 수 있다.

제시한 방법론의 검증을 위해 모의실험용 자료와 실제 시계열자료를 이용하였다. 실험결과 기존 연구에 비해 본 논문의 방법론에 의한 ARMA 모형화의 정확도가 높은 것으로 나타났다. 향후의 연구과제로는 인공신경망의 결합을 위해 다른 연산자의 이용이 가능할 것으로 사료되며 연산자간의 비교도 흥미있는 연구가 될 것이다.

[표-8] 실제 데이터에 대한 ARMA 모형화 결과

데이터	실제 ARMA 모형	제안된 ARMA 모형
Chemical Process Concentration[4]	ARMA(1,1)	ARMA(1,1)=0.8563 ARMA(1,4)=0.0538 ARMA(1,2)=0.0401
IBM Common Stock Closing Prices[4]	AR(1)	AR(1)=0.7642 AR(2)=0.1478 AR(4)=0.0406
Chemical Process Temperature Per Every Minute[4]	AR(2)	AR(2)=0.7972 AR(3)=0.0689 AR(5)=0.0553
Chemical Process Viscosity Per Every Hour[4]	AR(1)	AR(1)=0.7273 AR(2)=0.1195 AR(5)=0.0765
Wolfer Sunspot Data[4]	AR(2)	AR(2)=0.2665 AR(3)=0.2326 AR(4)=0.2125
Univ. of Wisconsin Hospital In-Patient Census[13]	ARMA(2,3)	ARMA(2,3)=0.4133 ARMA(3,2)=0.1693 ARMA(3,4)=0.1323
US Consumer Price Index[13]	ARMA(4,3)	ARMA(4,3)=0.9375 ARMA(2,1)=0.0441 ARMA(2,4)=0.0137
US Wholesale Price Index[13]	AR(1)	AR(1)=0.7112 AR(2)=0.0871 ARMA(1,5)=0.0532
Quarterly Gross National Product of the US[13]	AR(2)	AR(2)=0.4574 AR(3)=0.1841 AR(5)=0.1823
Quarterly Expenditures on Producer's Durables in the US[13]	AR(1)	AR(1)=0.5186 AR(2)=0.2636 AR(5)=0.1229

참고문헌

- [1] 오상봉, 이진창, "인공신경망 기초 의사결정트리분류기에 의한 시계열모형화에 관한 연구", 한국시물레이션학회 논문지, 제5권 제1호, 1996, pp.1-12.
- [2] Akaike, H. "A New Look at the Statistical Model Identification", IEEE Transactions on Automatic Control, vol.19, 1974, pp.716-723.
- [3] Beguin, J.M., C. Gourieroux and A. Monfort, "Identification of a Mixed Autoregressive Moving Average Process: The Corner Method", in Anderson, D.D. Ed., Time Series, North-Holland, Amsterdam, 1980, pp.423-436.
- [4] Box, G.E.P. and G.M. Jenkins, Time Series Analysis-Forecasting and Control, Holden-Day, San Francisco, 1970.
- [5] Cho, S. and Kim, J.H., "Combining Multiple Neural Networks by Fuzzy Integral for Robust Classification", IEEE Trans. Syst. Man Cybern., vol.25, 1995, pp.380-384.
- [6] Gray, H.L., A.G. Kelly and D.D. McIntire, "A New Approach to ARMA Modelling", Communications in Statistics, vol.B7, 1978, pp.1-77.
- [7] Hamilton, D.C. and D.G. Watts, "Interpreting Partial Autocorrelation Functions of Seasonal Time Series Models", Biometrika, vol.65, 1978, pp.135-140.
- [8] Jhee, W.C., K.C. Lee and J.K. Lee, "A Neural Network Approach for the Identification of the Box-Jenkins Model", Network, vol.3, 1992, pp.323-329.
- [9] Lee, J.K. and W.C. Jhee, "A Two-Stage Neural Network Approach for ARMA Model Identification with ESACF", Decision Support Systems, vol.11, 1994, pp.461-479.
- [10] Lee, K.C. and S.B. Oh, "An Intelligent Approach to Time Series Identification by a Neural Network-Driven Decision Tree Classifier", Decision Support Systems, vol.17, 1996, pp.183-197: **Highest Quality Rating by ANBAR Electronic Intelligence.**
- [11] Lee, K.C. and S.J. Park, "A Knowledge-Based Fuzzy Decision Tree Classifier for Time Series Modeling", Fuzzy Sets and Systems,

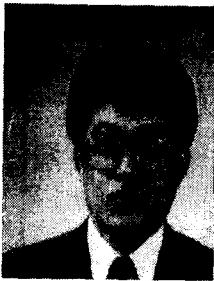
- vol.33, 1989, pp.1-18.
- [12] Lee, K.C. and S.J. Park, "PRTSM: Pattern Recognition-Based Time Series Modeler", Computer Science in Economics and Management, vol.2, 1989, pp.239-254.
- [13] Pandit, S.M. and S.M. Wu, Time Series and System Analysis with Applications, John Wiley and Sons, 1983.
- [14] Parzen, E. "ARMA Models for Time Series Analysis and Forecasting", Journal of Forecasting, vol.1, 1982, pp.67-82.
- [15] Rumelhart, D.E., G.E. Hinton and R.J. Williams, "Learning Internal Representations by Error Backpropagation", in Rumelhart, D.E. and J.L. McClelland, Eds., Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. I: Foundations, MIT Press, 1986.
- [16] Shafer, Glenn, A Mathematical Theory of Evidence, Princeton University Press, 1976, p.297.
- [17] Tsay, R.S. and G. C. Tiao, "Consistent Estimates of AR Parameters and ESACF for Stationary and Nonstationary ARMA Models", Journal of American Statistical Association, vol.79, 1984, pp.84-96.
- [18] Tucker, W.T. "On the Pade Table and Its Relationship to the R and S Arrays and ARMA Modelling", Communications in Statistics, vol.A11, 1982, pp.1335-1379.
- [19] Woodward, W.A. and H.L. Gray, "On the Relationship between S-Array and the Box-Jenkins Method of ARMA Model Identification", Journal of American Statistical Association, vol.76, 1981, pp.579-587.
- [20] Xu, L., Krzyzak, A. and Suen, C.Y., "Methods of Combining Multiple Classifiers and their Application to Handwriting Recognition", IEEE Trans. Syst. Man Cybern., vol.22, 1992, pp.688-704.
- [21] Yager, R.R., "Element Selection from a Fuzzy Subset using the Fuzzy Integral" IEEE Trans. Syst. Man Cybern., vol.23, 1993, pp.467-477.

Combining Multiple Neural Networks by Dempster's Rule of Combination for ARMA Model Identification

Sang-Bong Oh

A new methodology for ARMA Model Identification is suggested to rectify the limitations inherent in the Neural Network-driven Decision Tree Classifier(NNDTC), that is, any misclassification made at the parent node in the NNDTC cannot be corrected at any of its children nodes. It consists of the three stages: 1) ESACF(Extended Sample AutoCorrelation Function) extraction from the given time series data, 2) Local problem solving by individual neural networks, 3) Conflict Resolution. The methodology presented in this paper differs from the previous studies combining multiple neural networks: 1) The previous studies solve the same problem using multiple neural networks and their solutions are combined by some operator; 2) In this paper, a global problem is decomposed into multiple local problems and a neural network solves each local problem; After solving local problems, the local solutions are combined by Dempster's rule of combination in order to get an global solution. Experimental results with a set of simulated test data and real time series data show that the proposed methodology can efficiently identify the time series patterns with high precision compared to the previous approaches.

◆ 저자소개 ◆



오상봉 (Oh, Sang-Bong)

1993.4-현재 : 대전대학교 정보통신시스템공학부
부교수

1979.3-1983.2 : 서울대학교 경제학과 경제학사

1983.3-1990.2 : KAIST 경영과학과 공학석사, 공학박사

1990.4- 1993.4: 한국통신 경영전산부장(선임연구원)

현재 대전시 시정정책자문위원 및 정보화촉진협의회
위원

한국조폐공사 정보화추진위원회 위원

연구분야 : 전문가시스템, 인공지능응용, 전자상거래, 지
식기반예측시스템