

■ 論 文 ■

AIC(Akaike's Information Criterion)을 이용한 교통량 예측 모형

Traffic Forecasting Model Selection of Artificial Neural Network
Using Akaike's Information Criterion

강 원 의
(한국건설기술연구원 수석연구원)

백 남 철
(한국건설기술연구원 선임연구원)

윤 혜 경
(한양대학교 컴퓨터공학 석사)

목 차

- I. 서론
- II. 교통량 예측을 위한 신경망 이론
- III. 모형 설계
 - 1. 대상 자료
 - 2. 모형선택 기준
- 3. 신경망 모형 구성
- IV. 모형 분석 결과
- V. 결론
- 참고문헌

Key Words : 신경망 모델선택, 교통량 예측, 오류 역전파 알고리즘(BP), AIC(Akaike's Information Criterion), 시계열 분석

요 약

최근 교통량 예측을 위한 인공 신경망(Artificial neural networks : ANNs) 구조와 학습방법에 대한 연구가 다양하게 시도되고 있다. 이것은 신경망이 유연한 비선형 모형(non-linear model)으로 강력한 패턴 인식 능력을 가지고 있기 때문이다. 그러나, 신경망은 비선형 모형이기 때문에 많은 매개변수(parameter)를 사용하게 되면서 과적합(overfitting) 문제에 부딪히게 된다.

본 논문에서는 이러한 교통량 예측을 위한 신경망 모형에서 과적합을 해소하기 위한 방안으로 매개변수에 대한 다양한 모형선택기준(model selection criterion)에 대한 적용성에 대해서 알아보았다. 특히, AIC계열을 중심으로 모형선택기준으로 선택된 모형이 과적합 경향을 해소하고 시간적 전이성을 보장할 수 있는지를 분석하는데 본 연구의 목적을 두고 있다.

교통량 자료를 신경망 모형에 적용하여 분석한 결과, 첫째 학습자료(in-sample) 모형선택기준에 의해 선택된 모형이 검증자료(out-of-sample)의 최적의 성능을 보장하지는 못한다는 결과를 얻었다. 즉, 본 연구에서 기존의 연구에서처럼, 학습자료(in-sample)의 최적 모형이 검증자료(out-of-sample)의 성능과 직접적인 관계가 없다는 것을 알 수 있었다.

둘째 모형선택기준의 안정성을 분석한 결과 AIC3, AICC, BIC는 안정적인 모형을 선택하는 기준으로서 의미가 있는 것으로 분석되었다. 하지만, AIC4의 경우는 최상의 모형과 편차가 큰 것으로 분석되었다.

시계열 자료 분석과 예측에 있어서 모형의 불확실성은 학습 자료와 검증 자료의 상관관계에 영향을 줄 수 있음에 비춰볼 때, 앞으로 보다 많은 자료에 대한 분석이 필요하다고 판단되며, 다른 시계열 자료에 대한 분석이 요구된다.

I. 서론

최근 교통량 예측을 위한 인공 신경망(Artificial neural networks : ANNs) 구조와 학습방법에 대한 연구가 다양하게 시도되고 있다. 이것은 신경망이 유연한 비선형 모형(non-linear model)으로 강력한 패턴 인식 능력을 가지고 있기 때문이다. 그러나, 신경망은 비선형 모형이기 때문에 많은 매개변수(parameter)를 사용하게 되면서 과적합(overfitting) 문제에 부딪히게 된다.

본 논문에서는 교통량 예측을 위한 신경망 모형에서 과적합을 해소하기 위한 방안으로 매개변수에 대한 다양한 모형선택기준(model selection criterion)에 대한 적용성에 대해서 알아보았다. 특히, AIC계열을 중심으로 모형선택기준으로 선택된 모형이 과적합 경향을 해소하고 시간적 전이성을 보장할 수 있는지를 검토해 보고자 한다.

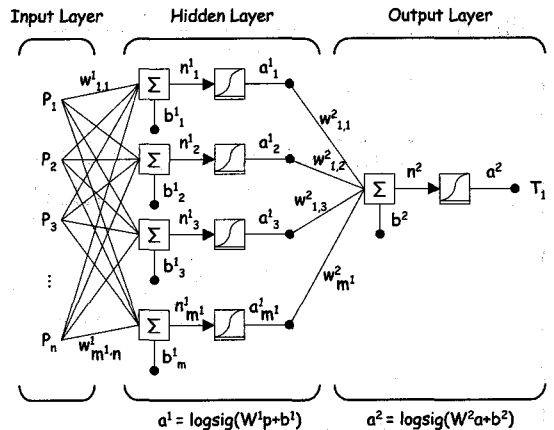
II. 교통량 예측을 위한 신경망 이론

신경망은 복잡한 구조를 가진 자료에서의 예측(prediction) 문제를 해결하기 위해서 사용되는 유연한 비선형 모형의 하나로 분류할 수 있다. 본 연구에서는 예측 어플리케이션에서 최근까지 가장 각광받는 패턴다임인 다층퍼셉트론 모형을 사용한다.

다층 퍼셉트론은 입력층과 출력층 사이에 하나 이상의 은닉층이 존재하는 신경망으로 네트워크는 입력층, 은닉층, 출력층 방향으로 연결되어 있으며, 각 층내의 연결과 출력층에서 입력층으로의 직접적인 연결은 존재하지 않는 전방향(feedforward) 네트워크(그림 1)이다.

학습 알고리즘은 출력층 오차를 이용하여 은닉층과 출력층간의 가중치를 변경하고, 출력층 오차를 은닉층에 역전파하여 입력층과 은닉층과의 연결강도를 변경하는 오류 역전파 알고리즘(error backpropagation algorithm)을 사용한다.

신경망은 유연한 비선형 모형(non-linear model)으로 강력한 패턴 인식 능력을 가지고 있다. 그러나, 비선형 모형이기 때문에 많은 매개변수(parameter)를 사용하게 되면서 과적합(overfitting) 문제에 부딪히는



(그림 1) 교통량 예측을 위한 단순화된 다층 퍼셉트론

한계성을 갖는다. 본 논문에서는 이러한 시계열 예측을 위한 신경망 모형에서 과적합을 해소하기 위한 방안으로 매개변수에 대한 다양한 모형선택기준(model selection criterion)에 대한 적용성에 대해 분석함으로써 신경망 이론이 갖는 한계성을 줄이고자 한다.

III. 모형 설계

1. 대상 자료

본 연구에서 사용하는 자료는 우리나라 일반국도의 교통량¹⁾을 사용하였다. 6개의 교통량패턴 그룹(TPGs: traffic pattern groups)²⁾에서 임의로 하나씩 지점을 선택하여 시계열 자료로 구성하여 사용하였다.

교통량 시계열 자료에 대한 표본 기간은 1997(1)부터 2001(52)까지 주별 평일 교통량 자료이다. 학습에 이용되는 In-sample 기간은 1997(1)부터 2000(52)까지이고, 모델에 이용되는 out-of-sample 기간은 2001(1)부터 2001(52)까지로 한다. 잠재적인 편의를 피하기 위해서 본 연구에서는 계절적인 영향을 받지 않는 특정일이 시계열 예측할 평일 자료와 중복되는 경우에는 예측할 평일의 전, 후일 자료를 참조하여 사용하였다.

2. 모형선택 기준

지금까지의 신경망 모형선택기준으로 주로 사용되었

1) 교통량이란 '도로의 한 지점(point), 차로(lane) 또는 구간(section)을 단위 시간 동안 통과하는 차량의 수'이며, 사용하는 목적에 따라 15분, 1시간, 일 교통량으로 집계한다.
 2) 교통량 특성이 유사한 상시조사지점들의 모임

〈표 1〉 모형 선택 기준

Criterion	Definition	Criterion	Definition
SSE	$\sum_{i=1}^T (y_i - \hat{y}_i)^2$	AICC	$\log\left(\frac{SSE}{T}\right) + \frac{2m}{(T-m-1)}$
AIC1	$\log\left(\frac{SSE}{T}\right) + \frac{2\log(m)}{T}$	BIC	$\log\left(\frac{SSE}{T}\right) + \frac{m\log(T)}{T}$
AIC2	$\log\left(\frac{SSE}{T}\right) + \frac{2\sqrt{m}}{T}$	RMSE	$\sqrt{\frac{1}{T} SSE}$
AIC3	$\log\left(\frac{SSE}{T}\right) + \frac{2m}{T}$	MAE	$\frac{1}{T} \sum y_i - \hat{y}_i $
AIC4	$\log\left(\frac{SSE}{T}\right) + \frac{2m^2}{T}$	MAPE	$\frac{1}{T} \sum \left \frac{y_i - \hat{y}_i}{y_i} \right $

〈표 2〉 학습 자료에서 선택된 최적 모형

Criterion	Best model (in-sample)					
	Spot #1	Spot #2	Spot #3	Spot #4	Spot #5	Spot #6
SSE	N(5,1)	N(5,2)	N(5,1)	N(4,1)	N(5,4)	N(5,4)
AIC1	N(5,1)	N(5,2)	N(4,1)	N(2,1)	N(5,4)	N(4,4)
AIC2	N(5,1)	N(5,2)	N(4,1)	N(2,1)	N(5,2)	N(4,4)
AIC3	N(2,1)	N(4,1)	N(4,1)	N(2,1)	N(1,1)	N(2,3)
AIC4	N(1,1)	N(1,1)	N(1,1)	N(2,1)	N(1,1)	N(1,1)
AICC	N(2,1)	N(4,1)	N(4,1)	N(2,1)	N(1,1)	N(3,2)
BIC	N(2,1)	N(4,1)	N(4,1)	N(2,1)	N(2,1)	N(2,1)
RMSE	N(5,1)	N(5,2)	N(4,1)	N(2,1)	N(5,4)	N(4,4)
MAE	N(2,5)	N(4,5)	N(5,4)	N(2,1)	N(2,1)	N(3,4)
MAPE	N(4,1)	N(4,5)	N(4,1)	N(4,1)	N(2,1)	N(3,2)

던 SSE는 신경망의 비선형의 특징 때문에 모형 예측치와 설계치의 과적합 문제를 발생시켰다. 이러한 SSE 사용시 문제점을 보완하기 위해서 Akaike(1969, 1973, 1974, 1978)는 SSE가 작아질수록 커지는 값을 매개변수로 하는 IC에 대한 값을 제외하였고 이에 대한 이론을 정립하였다. 현재는 AIC 기준 이후에 BIC의 기준이 정립되어 있는 상태이고, 본 연구에서는 〈표 1〉에서 정리한 모형선택기준을 적용하였다.

3. 신경망 모형 구성

입력 노드 수가 p 이고 은닉 노드수가 k 인 신경망에서 전체 매개 변수(total parameter)의 수(m)는 $m = k(p+2) + 1$ 이다.

각 입력 패턴에서 입력 노드와 은닉 노드수는 1에서 5까지 다양하게 구성하여 25개의 다른 신경망 모형을 구성하였다.

IV. 모형 분석 결과

시계열 교통량 자료에 대한 학습자료의 모형선택 결

〈표 3〉 검증 자료에서 선택된 최적 성능 모형

Criterion	Best model (out-of-sample)					
	Spot #1	Spot #2	Spot #3	Spot #4	Spot #5	Spot #6
RMSE	N(2,1)	N(3,5)	N(2,1)	N(2,3)	N(5,5)	N(4,4)
MAE	N(3,1)	N(3,3)	N(3,1)	N(3,5)	N(2,5)	N(1,1)
MAPE	N(3,1)	N(3,5)	N(3,1)	N(3,4)	N(2,5)	N(4,2)

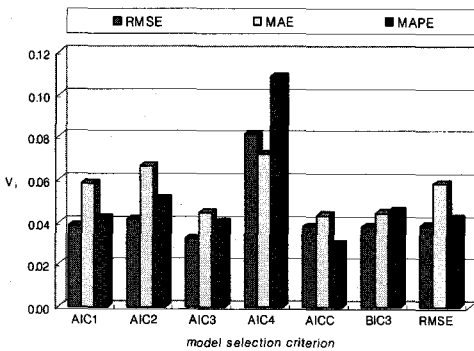
과와 검증자료의 성능 평가 결과는 각각 다음의 〈표 2〉와 〈표 3〉에 해당한다. 각 모형선택기준에 따라서 서로 상이한 모형을 선택하고 있음을 알 수 있다.

다음의 〈표 4〉는 상시조사지점 Spot #1에 대한 모형 선택 결과를 나타낸 것이다. 표를 통해 알 수 있듯이 AIC1-AIC2는 N(5,1) 모형을 선택했고 AIC3, AICC, BIC는 N(2,1) 모형을 선택하였으며, AIC4는 N(1,1) 모형을 선택하였다. 마찬가지로 분석한 모든 지점에서 모형선택기준에 의해서 선택된 모형이 서로 일치하는 경우도 있지만, 대개는 서로 다른 모형을 최상의 모형으로 선택한다. 검증자료(out-of-sample)의 성능 평가 기준별 결과를 보면, 학습자료(in-sample)의 모형선택기준에 의해서 선택된 모형이 일관된 결과를 보이지 않는다.

〈그림 2〉는 RMSE, MAE, MAPE의 3가지 성능 평가 기준에 대해서 모형선택기준의 안정성을 분석한 것이다. 모든 경우에 AIC3, BIC, AICC에 의해 선택

〈표 4〉 학습 자료와 검증자료에서 선택된 모형 비교(Spot #1)

p	k	m	in-sample							out-of-sample		
			RMSE ×100	AIC1	AIC2	AIC3	AIC4	AICC	BIC	RMSE ×100	MAE ×100	MAPE ×100
1	1	4	1.4545	-8.4477	-8.4418	-8.4225	-8.3172	-8.4216	-8.3584	1.9616	1.2352	10.1540
1	2	7	1.4556	-8.4408	-8.4341	-8.3922	-7.9884	-8.3895	-8.2799	1.9622	1.2364	10.1602
1	3	10	1.4561	-8.4367	-8.4284	-8.3627	-7.4973	-8.3573	-8.2022	1.9618	1.2353	10.1578
1	4	13	1.4566	-8.4334	-8.4234	-8.3331	-6.8331	-8.3241	-8.1245	1.9612	1.2339	10.1534
1	5	16	1.4567	-8.4314	-8.4196	-8.3042	-5.9965	-8.2905	-8.0475	1.9586	1.2291	10.1345
2	1	5	1.4187	-8.4953	-8.4892	-8.4625	-8.2693	-8.4611	-8.3820	1.5887	1.0555	7.7331
2	2	9	1.4203	-8.4873	-8.4796	-8.4216	-7.7259	-8.4172	-8.2767	1.5923	1.0621	7.7755
2	3	13	1.4211	-8.4827	-8.4726	-8.3818	-6.8746	-8.3727	-8.1725	1.5923	1.0634	7.7840
2	4	17	1.4216	-8.4794	-8.4669	-8.3425	-5.7145	-8.3269	-8.0688	1.5912	1.0630	7.7832
2	5	21	1.4218	-8.4771	-8.4622	-8.3036	-4.2456	-8.2795	-7.9655	1.5892	1.0609	7.7733
3	1	6	1.4197	-8.4920	-8.4856	-8.4512	-8.1599	-8.4491	-8.3542	1.5908	1.0409	7.5818
3	2	11	1.4214	-8.4838	-8.4748	-8.4002	-7.3323	-8.3936	-8.2225	1.5950	1.0490	7.6385
3	3	16	1.4223	-8.4789	-8.4670	-8.3505	-6.0204	-8.3365	-8.0920	1.5960	1.0516	7.6584
3	4	21	1.4228	-8.4755	-8.4606	-8.3012	-4.2236	-8.2768	-7.9620	1.5954	1.0523	7.6657
3	5	26	1.4230	-8.4732	-8.4553	-8.2524	-1.9417	-8.2143	-7.8324	1.5942	1.0510	7.6629
4	1	7	1.4082	-8.5068	-8.4999	-8.4575	-8.0477	-8.4547	-8.3440	1.6434	1.0711	7.6369
4	2	13	1.4104	-8.4975	-8.4874	-8.3957	-6.8738	-8.3864	-8.1850	1.6403	1.0757	7.6586
4	3	19	1.4111	-8.4928	-8.4790	-8.3362	-4.9996	-8.3161	-8.0282	1.6379	1.0754	7.6550
4	4	25	1.4119	-8.4891	-8.4717	-8.2766	-2.4229	-8.2411	-7.8713	1.6342	1.0746	7.6457
4	5	31	1.4121	-8.4868	-8.4659	-8.2178	0.8554	-8.1619	-7.7153	1.6331	1.0737	7.6440
5	1	8	1.4027	-8.5132	-8.5058	-8.4551	-7.9061	-8.4515	-8.3250	1.7266	1.1237	7.9697
5	2	15	1.4053	-8.5033	-8.4919	-8.3828	-6.3240	-8.3703	-8.1388	1.7186	1.1245	7.9592
5	3	22	1.4063	-8.4981	-8.4824	-8.3127	-3.7833	-8.2853	-7.9549	1.7122	1.1215	7.9327
5	4	29	1.4068	-8.4947	-8.4749	-8.2434	-0.2826	-8.1944	-7.7717	1.7100	1.1208	7.9284
5	5	36	1.4074	-8.4917	-8.4680	-8.1739	4.1790	-8.0957	-7.5884	1.7065	1.1185	7.9124



〈그림 2〉 모형선택기준의 안정성 평가 결과

된 모형들은 최적의 모형과 비교하여 안정적인 성능을 발휘하는 것으로 평가되었다. 반면에 AIC4의 경우는 최상의 모형과 편차가 큰 것으로 판측되었다.

V. 결론

최근 교통량 예측 분야에서 신경망 분야에 대한 관

심이 높아지고 있다. 하지만, 신경망은 일반적으로 다량의 매개변수(parameter)로 평가되어지기 때문에, 과적합하는 경향이 있으며 이러한 경향을 개선하기 위해서 본 연구에서는 모형선택기준을 적용하는 방법에 대해서 유효성을 실험적으로 검토하였다. 실험 자료로는 시계열 교통량 자료를 이용하여 분석하였다.

분석결과 학습자료(in-sample) 모형선택기준에 의해 선택된 모형이 검증자료(out-of-sample)의 최적의 성능을 보장하지는 못했다. 즉, 본 연구에서 기존의 연구에서처럼, 학습자료의 최적 모형이 검증자료의 성능과 직접적인 관계가 없다는 것을 알 수 있었다. 따라서 학습자료와 검증자료에서 공통적으로 모형성능을 만족시키는 정보평가 기준이 필요하다. 이에 본 연구에서는 모형선택기준의 안정성을 분석한 결과 AIC3, AICC, BIC는 안정적인 모형을 선택하는 기준으로서 의미가 있는 것으로 분석되었다. 앞으로 보다 많은 자료에 대한 분석이 필요하며, 다른 시계열 자료에 대한 분석이 요구된다.

참고문헌

1. 백남철(2002), 베이지안 샘플링과 신경망관별함수를 이용한 교통량 패턴그룹기반의 AADT추정(Traffic Pattern Groups-based estimation of AADT using Bayesian Sampling and Neural Network Discriminant Functions), 서울대학교 대학원 박사학위논문, pp.57~77.
2. 윤혜경(2002), AIC를 이용한 신경망모델 선정에 관한 연구, 한양대학교, 석사학위논문.
3. Akaike, H.(1969), Fitting autoregressive models for predictions. Annals of the Institute of Statistical Mathematics 21, pp.243~247.
4. Akaike, H(1973), 2nd International Symposium on Information Theory, chapter Information theory and an extension of maximum likelihood principle, pp 267-281. Budapest: Akademia Kiado
5. Akaike, H(1974), A new look at the statistical model identification IEEE Transactions on Automatic Control AC-19, pp.716~723.
6. Akaike, H(1978), Abayesian analysis of the minimum aic procedure, Annals of the Insititute of mathematical Statistics A(30), pp.9~14.
7. Bishop, C.M.(1995), Neural network for pattern recognition, pp 116-140, pp.371~380.
8. Fausett, L.(1994), Fundamental of neural networks, pp.11~21, pp.289~316.
9. Martin, T.H., Howard, B.D. and Mark, B. (1996), Neural network design, chap. 11, pp.1~25.

✉ 주 작 성 자 : 강원의
 ✉ 논문투고일 : 2004. 10. 13
 논문심사일 : 2004. 10. 29 (1차)
 심사판정일 : 2004. 10. 29
 ✉ 반론접수기한 : 2005. 4. 30