

■ 論 文 ■

신경망 이론을 이용한 통행발생 모형연구 (선형/비선형 회귀모형과의 비교)

Trip Generation Model Using Backpropagation Neural Networks
in Comparison with linear/nonlinear Regression Analysis

장 수 은

김 대 현

임 강 원

(서울대학교 환경대학원 석사과정) (한국도로공사 도로연구소 책임연구원) (서울대학교 환경대학원 교수)

목 차

- I. 서론
- II. 기존 통행발생모형
 - 1. 선형 회귀모형
 - 2. 지수모형(로그 - 선형모형)
 - 3. 비선형 회귀모형
- III. 신경망이론을 이용한 통행발생 모형
 - 1. 신경망 이론의 개요
- 2. 신경망 이론의 적용
- IV. 모형의 비교검정
 - 1. 분석자료 및 연구방법
 - 2. 비교평가
 - 3. 신경망 모형의 교사치 및 추정치 분석
- V. 결론
참고문헌

요 약

본 연구의 목적은 기존의 대표적 통행발생모형인 회귀모형과 신경망 이론에 의한 통행발생모형을 비교·분석하여 통행발생모형에 대한 새로운 방법을 제시하고자 하는 것이다. 이를 위해 모형의 검정력과 안정성을 현재적 설명력과 장래 예측력의 결합으로 전제하고, 시나리오에 따른 모형의 검정력 변화를 통한 안정성 평가를 수행하였다.

연구결과 역전파 신경망 모형(Backpropagation Neural Networks)은 회귀모형의 검정력과 안정성을 상회하는 우수한 결과를 보여 주었으며, 이는 향후 통행발생 모형으로 역전파 신경망 모형의 적용 가능성을 의미하는 것으로 해석된다. 특히 복잡해진 교통현상과 다양한 수집자료를 고려할 때 교통분야에서의 신경망 모형의 적용은 더욱 확대될 전망이다.

1. 서론

교통수요 분석에서 전통적으로 가장 많이 사용되어 오면서 도시교통 모형의 근간을 이루고 있는 방법은 4단계 추정법이다. 이 접근법은 그 동안 다양한 도시교통 여건에 적용되었으며 대부분의 교통수요 예측에 표준적 과정으로 자리잡고 있다. 4단계 추정법이 다른 접근법과 구별되는 특징은 수요예측의 순차성에 있다. 즉, 『통행발생, 통행배분, 수단선택, 노선배정』의 개별적 과정을 거친 뒤, 모든 단계의 균형점에서 완결된 형태로 종결된다. 이 과정에서 개별 단계는 선행단계의 추정결과에 근거해 모형을 수립하므로 선행단계의 중요성은 매우 크다고 할 수 있다. 특히 통행발생은 4단계 수요모형의 첫 단계로서, 이 단계의 추정결과가 전체 수요예측에 크게 영향을 미치므로 더욱 중요한 의미를 지닌다. 따라서 통행발생 단계에서 무엇보다도 필요한 것은 추정의 정확성이라 할 수 있다.

일반적으로 통행발생은 계획대상 도시지역의 장래 사회경제적 지표에 대한 정량적 예측을 기반으로 각 존별 통행의 발생량을 예측하는 것이다. 통행발생량은 통행자의 속성(직업, 연령, 성별 등)과 통행목적의 결합으로 복잡한 발생특성을 띠는 것으로 알려져 있다. 여기에 통행자의 여건변화에 따른 주관적 의사결정 과정을 수반하게 되므로 복잡한 함수관계를 나타내게 된다.

현재 통행발생 모형으로 널리 사용되고 있는 것은 선형 회귀모형이다. 선형 회귀모형은 오차에 대한 무리한 가정과 각종 사회경제적 지표와 통행발생량의 관계가 모수(parameter)에 선형(linear in the parameters)임을 전제로 한다. 경직된 가정은 현실을 왜곡하게 되고 그 결과 자료에 따라 모형의 검정력에 심한 변동을 보여 왔다. 이러한 결함에도 불구하고 회귀모형이 널리 이용된 이유는 적용의 간편성 뿐만 아니라, 결과해석의 용이성 때문이었다.

반면, 최근 교통을 포함한 거의 모든 학문분야에 널리 이용되고 있는 신경망 이론은 복잡하고 부정확한 데이터나 특별한 사전정보가 없는 경우를 모형화 하는데 우수한 능력을 발휘하는 것으로 인식되고 있다. 특히 역전파(Backpropagation) 모형으로 불리는 역전파 학습 알고리즘을 이용한 다층전파 네트워크의 경우 적용이 간편하며, 여러 연구에서 모형의 우수성이 확인되고 있다(Mark Dougherty, 1995).

통행발생모형 역시, 각종 사회경제적 변수와 통행발생의 복잡한 함수관계를 모형화 하는 것이므로 신경망 이론의 적용은 유용한 것이라 하겠다.

이에 본 연구는 기존의 대표적인 통행발생모형인 회귀모형과 신경망 이론에 의한 통행발생모형과의 결과를 비교·분석하고, 이를 토대로 통행발생모형에 대한 새로운 접근법을 제시하고자 한다. 이를 위해 모형의 검정력과 안정성을 현재적 설명력과 장래 예측력의 결합으로 전제하고, 시나리오에 따른 모형의 검정력 변화를 통한 안정성 평가를 수행한다. 또한 지금까지의 연구에서 전제되었던 완전정보의 현실과리성을 극복하기 위해 제한된 정보로부터 추가정보에 따른 회귀모형과 역전파 신경망(Backpropagation Neural Networks)모형의 검정력의 변화를 비교·분석한다. 이를 토대로 모형간 안정성의 차이 및 변화를 밝히고, 기존 통행발생 모형을 대체할 새로운 접근법을 제시하고자 한다.

II. 기존 통행발생 모형

1. 선형 회귀모형

선형 회귀모형은 갈튼(1822~1911)의 연구 이후 여러 학문 영역에 적용되는 모형으로서 교통 수요분석에서 가장 광범위하게 적용되고 있다.

$$Y = X\beta + \epsilon, \epsilon \sim N(0, I\sigma^2) \quad (1)$$

Y : 통행발생 벡터

X : 설명변수 벡터

ϵ : 오차항

β : 파라메타

σ^2 : 분산

N : 정규분포

I : 항등행렬

$$\begin{aligned} \text{Cov}(\epsilon_i, \epsilon_j) &= 0, (i \neq j) \\ &= \sigma^2, (i = j) \end{aligned} \quad (2)$$

Cov : 공분산

식(1)에서 $\text{Var}(\epsilon) = I\sigma^2$ 으로 오차들은 각각 동일한

분산 $\text{Var}(\epsilon_i) = \sigma^2$ 를 갖고 오차들끼리는 서로 독립임을 가정한다. 선형 회귀모형은 식(2)와 같은 오차에 대한 가정뿐만 아니라 설명변수와 종속변수의 관계를 “모수에 선형(linear in the parameters)”으로 설정하고 있다. 이러한 무리한 가정으로 인해 자료에 따른 모형의 검정력에 심함 변동을 내재함이 알려져 있다.

2. 지수모형(로그-선형모형)

지수모형은 1960~1970년대 중력모형(gravity model)의 정산(calibration)과 추상수단분담모형(abstract mode split model)의 파라메타 추정에 적용되기 시작하여 현재 여러 교통 수요분석에 광범위하게 적용되고 있는 모형이다(Oum, 1989). 이 모형은 식(3)과 같은 Cobb - Douglas 함수형태를 이룬다.

$$Y = X^\beta \epsilon, \quad \epsilon \sim N(0, I\sigma^2) \quad (3)$$

식(3)에 자연대수를 취하면 선형 회귀모형과 동일한 형태를 이루게 된다(식(4)).

$$\ln Y = \ln X^\beta + \ln \epsilon, \quad \ln \epsilon \sim N(0, \sigma^2) \quad (4)$$

지수모형의 구조는 추정된 파라메타가 개별 변수의 수요탄력성 값이 되어 정책적 적용이 간편하며 제한된 범위의 비선형 관계를 모형화하는 장점이 있다.

3. 비선형 회귀모형

비선형 회귀모형은 선형 회귀모형의 현실과의 괴리를 극복하고자 제한된 것이다. 통행발생은 일반적으로 비선형 발생특성을 나타내므로 이 접근법은 기존 방법에 비해 진일보한 방법으로 인식되고 있다. 이 모형의 범주에는 비선형 최소제곱법(nonlinear Ordinary Least Square), 2단계 최소제곱법(Two Stage Least Square) 등의 방법이 있지만, 이 방법은 선형 최소제곱추정의 구간분할을 통한 불연속 선형근사(piecewise linear approximation)라는 한계를 지니고 있다.

반면 본 연구에 적용된 Box-Cox 회귀모형은 Box-Cox변환(1964)을 모형내에 포함한 것으로써 최적 파라메타 결정을 통한 비선형 관계를 모형화한다. 이와 같은 유연함으로 인해 그동안 교통수요 분석(Gaudry

and Wills, 1978, Oum, 1989)과 교통수단 선택 연구(Mandel B. et al, 1994, Picard G. and M. Gaudry, 1998) 등에 광범위하게 적용된 모형이다.

$$g(Y) = f(X, \beta) + \epsilon; \quad (5)$$

식(5)에서 비선형을 가정하는 이유는 함수 f의 비선형성과 독립변수 X와 오차 ϵ 의 상관관계에 기인한다. 식(5)를 식(6)과 같이 변환(Box-Cox변환)하면 식(7)의 비선형 회귀모형이 유도된다.

$$\begin{aligned} \lambda \neq 0 \text{ 인 경우} \quad & X^{(\lambda)} = \frac{X^{(\lambda)} - 1}{\lambda} \\ \lambda = 0 \text{ 인 경우} \quad & X^{(\lambda)} = \ln X \end{aligned} \quad (6)$$

$$Y = \beta X^{(\lambda)} + \alpha Z + \epsilon \quad (7)$$

X : 비선형 독립변수
Z : 선형 독립변수
 λ : 비선형 파라메타

이 모형은 식(7)에서 보듯 통행발생에 영향을 미치는 변수를 선형과 비선형 변수로 구분해 개별 변수의 효과를 모형내에 내재하는 장점을 지니고 있다. 또한 식(6)에서 $\lambda = 1$ 인 경우는 선형 회귀모형이 되며, $\lambda = 0$ 인 경우는 지수모형(로그-선형 모형)이 됨을 알 수 있다. 결국 Box-Cox 모형은 회귀모형의 일반적 형태라 할 수 있다.

III. 신경망 이론을 이용한 통행발생 모형

1. 신경망 이론의 개요

인공 신경망(Artificial Neural Networks)은 흔히 신경망(Neural Networks)라 불리며, 이는 생물학적인 학습과정에 기초를 두고 있는 컴퓨터의 응용이다. 인공 신경망은 McCulloch와 Pitts(1943)의 연구를 시초로 하여, Rumelhart와 MacClelland(1986) 등에 의한 연구로 인해 인공 신경망, 특히 현재 가장 널리 사용되고 있는 역전파 신경망 모형(Backpropagation Neural Networks)에 대한 관심이 전세계적으로 고조되었다. 신경망 이론은 복잡하거나 부정확

한 데이터로부터 추론할 때 우수한 능력을 발휘하며, 인간이 직접 또는 기존의 컴퓨터 모델을 이용하여서는 해석하기 어려운 문제에 유용하게 사용되어 왔다. 현재 신경망 이론은 수학, 컴퓨터 공학, 토목 공학, 교통, 의학, 생물학, 심리학 등 거의 모든 학문분야에서 널리 이용되고 있으며 새로운 모형의 개발과 더불어 신경망 이론에 대한 인식은 더욱 증가될 전망이다.

현재 다양한 종류의 신경망 모델이 개발, 이용되고 있으나 흔히 역전파 모형(Backpropagation)으로 불리는 역전파 학습 알고리즘을 이용한 다층전파 네트워크가 가장 널리 이용되는 신경망 모형 중의 하나이다. 이는 역전파 모형이 사용하기 쉽고 여러 분야에 적용 가능하다는 장점 외에 지금까지의 많은 연구 결과 모형의 우수성이 인정되었기 때문이다(Mark Dougherty, 1995).

2. 신경망 이론의 통행발생모형 적용

역전파 모형(Backpropagation)은 사용하기에 앞서 네트워크 구성 및 학습변수값, 학습방식 등을 결정해야 한다. 따라서, 본 연구는 다음의 순서로 모형을 구성한다. 우선 분석 네트워크는 <그림 1>과 같이 입력변수(input unit)와 은닉층 뉴런(hidden neuron), 출력변수(output unit)로 구성한다.

여기서 최적 은닉층 뉴런의 수가 결정되어야 하는데, 이에 관한 일반적 방법은 없으며 분석대상 자료 특성에 가장 부합하는 최적 뉴런의 수는 경험적 오류수정 과정을 통해 결정된다. 본 연구는 여러 시행착오(trial

and error)를 거쳐 5개의 은닉층을 설정하였다. 따라서 연구 네트워크는 3개의 입력변수와 5개의 은닉층 뉴런, 하나의 출력변수로 구성된다.

구성된 네트워크의 각 층(layer)간 정보의 전이는 활동함수(activation function)를 거치며 이루어진다. 본 연구에서 입력변수와 은닉층사이에는 -1과 1의 범위의 값을 갖는 Hyperbolic Tangent함수를, 은닉층과 출력변수사이에는 Piecewise-linear 함수를 각각 적용하였다(<그림 2>). 왼쪽의 "Hyperbolic Tangent" 함수는 통행발생의 비선형 관계를 사상(mapping)하기 위함이고, 오른쪽의 "Piecewise-linear" 함수는 학습결과를 목표단위(Target)로 연결하기 위함이다.

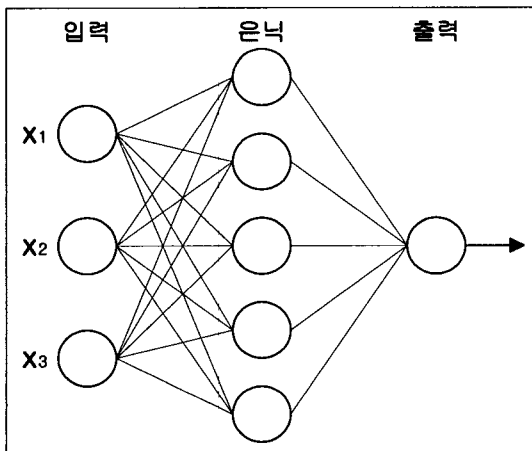
학습(training) 알고리즘은 뉴턴-랩슨 알고리즘을 수정한 "Levenberg-Marquardt 알고리즘"을 적용하였다. 이 방법은 헤시안 행렬을 계산할 필요 없이 2계 조건을 근사추정할 수 있도록 제안된 것으로 이 알고리즘에 의하면 헤시안 행렬과 gradient는 (8)과 같이 표현된다. 이처럼 헤시안행렬(H)이 야코비안행렬(J)로만 구성되므로, 신속하게 최적값에 수렴하는 장점을 지니고 있다.

$$H = J^T J$$

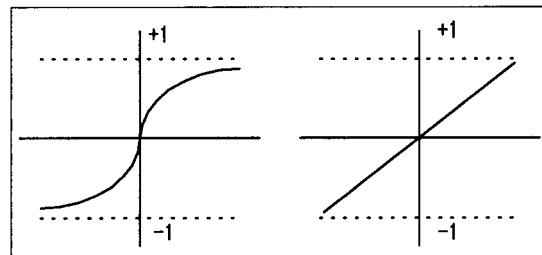
$$g = J^T e \tag{8}$$

- H : 헤시안 행렬
- J : 야코비안 행렬
- e : 오차벡터

최적 연결강도(weight) 조정은 "일반화 증감법 (Generalized Delta Rule)"을 적용(Dayhoff, 1986) 하였다. 식(9)에서 적률상수(MC; Momentum Constant)와 학습률(LR; Learning Rate)은 모두 최적



<그림 1> 네트워크 구조



<그림 2> 활동함수

연결강도(weight)을 얻기 위한 파라메타로, 적률상수(MC)는 국지적 최소값이 발생하는 것을 조정하기 위해 가중치의 방향을 일정하게 유지시키는 역할을 하며, 학습률(LR)은 최적 연결강도를 얻기 위한 학습변수값을 나타낸다.

$$w_{ij}(n+1) = w_{ij}(n) + [\Delta w_{ij}(n) + \alpha \Delta w_{ij}(n-1)]$$

$$\Delta w_{ij} = \eta \frac{\partial(\text{error}_i)}{\partial(w_{ij})} \text{out}_i \quad (9)$$

- $w_{ij}(n)$: n 번째 반복시 연결강도
- $\Delta w_{ij}(n)$: n 번째 반복시 연결강도조정분
- α : MC
- η : LR

은닉층의 수를 결정하는 것과 마찬가지로 최적의 MC와 LR을 결정하는 일반적인 방법은 없으며, 이 또한 경험적 조정과정을 거치게 된다. 많은 연구에서 최적 MC와 LR은 0.4~0.6사이에서 결정되고 있으므로 본 연구에서도 0.4~0.6사이에서 MC와 LR의 초기값을 무작위로 부여한 뒤, MSE(Mean Squared Error)를 최소로 하는 미세한 반복계산 조정과정(± 0.08)을 통해 최적 연결강도에 수렴하는 방식을 택했다.

마지막으로 자료의 과다학습(overfit) 문제를 방지하기 위해 자료를 구분 적용하는 교차검증법(Cross-validation)법을 적용하였다. 모든 자료에는 일반적으로 이상점(outlier)이 존재하며 이는 관측 오차일수도 있고, 입력과정의 실수일수도 있다. 그런데, 신경망의 학습과정에서 오차를 과다하게 줄이게 되면, 이상점까지 모형에 내재시켜 예측력의 감소를 가져 온다. 모형을 추정한다는 것은 자료의 정확한 적합(fitting exactly)보다 자료의 구조적 경향(structural trend)를 추정하는데 있으므로 학습자료와 검정자료의 오차 균형을 교차검정법을 통해 유지하는 방법을 택했다.

IV. 모형의 비교검정

1. 분석자료 및 연구방법

1) 분석자료

본 연구의 분석자료는 교통개발연구원의 『제 1차

전국 물류현황 조사 - 지역간 화물 물동량 조사』의 최종보고서에 제시된 15개 대존¹⁾의 1995년 11월 18일~11월 20일까지의 3일간 물동량 자료이다. 물류현황 자료에는 한국 표준 산업 분류에 의한 33개 품목의 물동량이 조사되었으나, 본 연구는 그 중 “경공업품”의 유출(Production) 물동량에 한정하여 분석하였다. 고려된 변수는 인구, GRP, 도소매업 종사자수, 서비스업 종사자수, 지역면적 등 모두 5가지이며, 상관분석을 통해 유의한 변수 3가지를 선택했다.

2) 자료의 검정

〈표 1〉에 제시된 자료의 관측점 수는 모두 15개로 중심극한정리(central limit theorem)를 만족하는 유효 표본수(25개)에 못 미친다. 이러한 현상은 대존을 대상으로 하는 국내 연구에서 언제나 발생하는 문제로써 선택된 횡단면(cross-section) 표본의 정규성 검정이 필요하며 검정결과 자료가 정규성 조건을 만족한다면 분석의 통계적 근거가 마련되는 것이고, 그렇지 않다면 3~4년 정도의 시계열 자료를 추가 조사한 통합자료(pooling data)를 구비하는 등의 조치가 필요하다. 이 경우 자기상관(autocorrelation)의 문제나 이분산 문제를(heteroscedasticity) 해결해야

〈표 1〉 연구에 적용된 자료

| 존 | 물동량 (톤) | 인구 (명) | GRP (백만원) | 지역면적 (m ²) |
|----|---------|------------|------------|------------------------|
| 1 | 220,426 | 10,231,217 | 62,166,311 | 14,920,596 |
| 2 | 124,955 | 3,814,325 | 17,077,158 | 6,057,118 |
| 3 | 586,483 | 2,449,420 | 10,238,870 | 4,344,092 |
| 4 | 65,463 | 2,308,188 | 13,219,694 | 3,342,845 |
| 5 | 29,760 | 1,257,636 | 5,776,577 | 2,241,349 |
| 6 | 149,396 | 1,272,121 | 5,532,871 | 2,463,181 |
| 7 | 157,216 | 7,649,741 | 45,842,819 | 11,284,785 |
| 8 | 56,752 | 1,466,238 | 6,547,890 | 2,547,572 |
| 9 | 54,110 | 1,396,728 | 8,074,995 | 2,577,648 |
| 10 | 36,864 | 1,766,854 | 9,838,622 | 3,252,831 |
| 11 | 76,065 | 1,902,044 | 8,877,967 | 2,987,713 |
| 12 | 38,106 | 2,066,842 | 14,179,990 | 3,270,944 |
| 13 | 64,620 | 2,676,312 | 17,116,379 | 3,997,866 |
| 14 | 146,185 | 3,845,622 | 30,185,017 | 6,477,781 |
| 15 | 762 | 505,438 | 2,407,492 | 1,206,653 |

1) 15개 존은 시·도 단위로 구분된 대존으로 1개 특별시와 5개 광역시 9개 도로 구성된다.

〈표 2〉 정규성 검정 결과

| 구 분 | 통 계 량 |
|---------------|---------|
| 평균(Mean) | 11.4078 |
| 표준편차(Std Dev) | 0.81738 |
| 왜도(Skewness) | 0.77416 |
| 첨도(Kurtosis) | 0.53975 |
| W : Normal | 0.94078 |
| Pr(W) | 0.41020 |

함은 물론이다. 본 연구는 SAS의 Shapiro-Wilks의 정규성 검정을 시행했으며, 주요 결과는 〈표 2〉와 같다.

왜도(Skewness)의 경우 완전 대칭상태인 "0"에 매우 가까우며 첨도(Kurtosis) 역시 정규분포를 만족하는 유효범위(-1≤첨도≤5)에 포함되어 있음을 알 수 있다. 또한 정규성 판단의 유의 확률 역시 비교적 양호하므로 본 자료는 정규분포에 근사하다는 판단을 내릴 수 있다.

3) 연구모형

화물통행 발생함수는 식(10)과 같이 구축하였다. 식에서 화물 물동량(Y)은 총물동량이 아닌 순물동량으로서 화물의 출발점에서 도착점까지의 이동을 하나의 유동으로 파악하는 개념을 적용했으며, 이동중 운송수단이 변할 경우 운송거리가 가장 긴 수단에 근거해 물동량을 파악했다.

$$Y = f(X_1, X_2, X_3) \tag{10}$$

- Y : 물동량
- X₁ : 인구
- X₂ : GRP
- X₃ : 지역면적

적용된 설명변수는 〈표 1〉에 제시된 인구(X₁), 지역내 총생산(X₂), 지역면적(X₃) 등이다.

4) 연구수행방법

본 연구의 수행방법은 불완전정보로부터 추가정보가 주어질 때 모형의 안정성 변화를 분석하는 것이다. 이를 위해 초기 정보를 난수발생에 의해 설정했으며, 추가정보의 확장 역시 임의추출법을 사용하였다. 구체적인 연구의 수행방법은 다음과 같다.

- ① 난수발생을 통한 11개 존 자료를 추출, 모형을 추정한다.
- ② 초기정보를 복원한 후, 정보의 양을 늘려가며 모형의 변화를 검정하고 완전정보에 접근한다.
- ③ 개별단계의 모수 및 각종 통계치의 변화를 분석한다.
- ④ 모형의 신뢰성과 안정성에 관한 종합분석을 수행한다.

5) 평가기준

모형간 안정성의 평가기준으로는 예측값(교사치), 결정계수(R²), 평균제곱오차(MSE; Mean Squared Error), 평균상대분산도(ARV; Average Relative Variance)를 적용한다. 예측값과 결정계수는 모형의 전반적인 검정력을 분석하는 것이며, 평균제곱오차(MSE)와 평균상대분산도(ARV)는 모형의 오차 및 분산정도를 측정하는 지표이다.

$$R^2 = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2} \tag{11}$$

- Y_i : 물동량 실측치
- \hat{Y}_i : 물동량 추정치

$$MSE = \frac{\sum(Y_i - \hat{Y}_i)^2}{N} \tag{12}$$

$$ARV = \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2} = \frac{1}{\sigma^2} \frac{1}{N} \sum(Y_i - \hat{Y}_i)^2 \tag{13}$$

- N : 관찰점 수
- σ^2 : 분산 추정치

결정계수는 모형의 적합도를 비교할 수 있는 장점이 있지만 각 모형이 구체적으로 어느 정도의 오차를 갖고 있는지를 보여주지 못한다는 한계점이 있다. 반면에 오차와 분산도는 모형의 적합도를 비교할 수는 없지만 각 모형이 포함하고 있는 오차를 보여준다. 따라서 모형간 성과도를 비교하고자 할 때 이 두 가지를 모두 고려하는 것이 바람직하다고 할 수 있다.

2. 비교평가

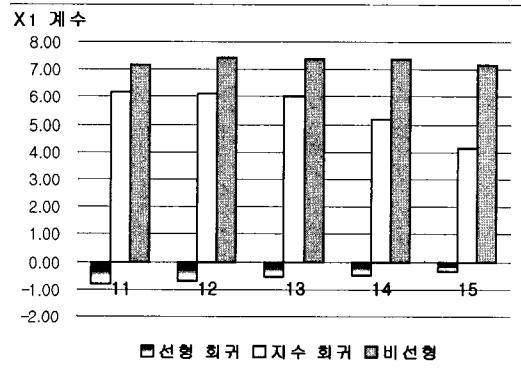
모형간 성과도의 차이 및 안정성 검정을 위하여 네

가지 모형을 비교·분석하였다. 회귀모형은 선형 회귀모형, 지수모형, 비선형 모형(Box-Cox모형)을 사용하였으며, 신경망 모형과 그 결과를 비교하였다. 선형 회귀모형과 지수모형의 추정에 이용된 프로그램은 SAS이며, 비선형 모형에는 LIMDEP, 신경망 모형에는 MATLAB을 각각 사용하였다.

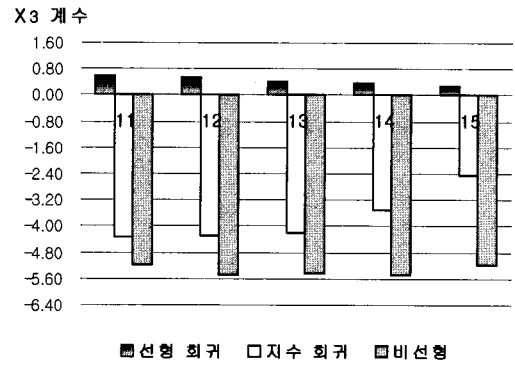
1) 회귀분석결과

회귀분석에서 부호의 상식성 여부는 모형의 성립 여부를 판단하는 가장 기본적인 기준이다. 초기분석결과 X_2 의 부호가 음으로 나타났다. GRP가 경공업품 발생량에 음의 영향을 미치는 것은 상식적으로 잘못된 관계이므로 이를 제외한 나머지 변수로 모형을 추정하였다. 최종 모형은 선형 회귀모형을 제외한 지수모형과 비선형 모형에서 모두 상식적 부호를 나타내었다. 즉 인구(X_1)는 양의 관계를 나타내고 지역면적(X_3)은 음의 관계를 보여주어, 경공업품에 대한 수요가 인구가 많은 도시부 지역에서 주로 발생하는 합당한 결과를 나타내었다. 모형간 파라메타를 살펴보면 비선형 모형이 가장 큰 수치를 나타내었고, 선형 회귀모형이 가장 작은 수치를 나타내는데 그 차이가 무려 14~15배에 달하고 있다. 이를 개별 설명변수별로 살펴보면 아래와 같다.

우선 X_1 (인구)의 경우 비선형 회귀모형은 추가정보에 따라 7.19~7.41 사이의 값을, 지수모형 4.11~6.17, 선형 회귀모형은 -0.81~-0.35 사이의 값을 나타내고 있다. 지수모형과 선형 회귀모형은 계수간 편차가 심해 모형의 신뢰성이 낮으며, 특히 선형 회귀모형의 경우 계수의 부호가 상식과 일치하지 않는 불합리한 결과를 나타내었다.



〈그림 3〉 X_1 계수의 변화추이



〈그림 4〉 X_3 계수의 변화추이

X_3 (지역면적)의 경우 비선형 모형이 -5.48~-5.17, 지수모형이 -4.33~-2.45, 선형 회귀모형이 0.26~0.59 사이의 값을 나타내 X_1 의 경우와 유사하였다. 이 경우에도 선형 회귀모형의 경우 부호가 상식과 일치하는 않는 불합리한 결과를 나타냈다.

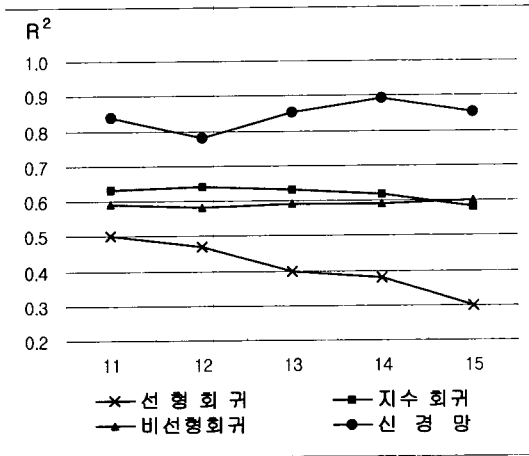
전반적으로 볼 때, Box-Cox 비선형 모형의 모수 추정치가 비교적 균일하며, 추가정보에 따른 편차가 가장 적은 것으로 나타났다. 지수모형은 로그변환을 통한 비율조정효과로 비교적 안정된 모수 추정치를 제시하고 있다. 반면, 선형 회귀모형은 변동폭과 편차가 크고 부호가 상식과 일치하지 않아 모형적용의 한계를 나타내었다.

2) 모형비교

모형의 전반적 검정력을 평가하는 지표로 널리 이용되는 것은 결정계수이다. 아래 〈표 3〉과 〈그림 5〉는 추가정보에 따른 결정계수의 변화추이를 나타낸 것이다. 결과에서 나타난 바와 같이 비선형 회귀모형과 지수모형의 검정력은 비교적 양호하며 신경망 모형의 검정력은 우수한 것으로 나타났다. 신경망 모형은 안정된 패턴으로 높은 수치를 나타내고 있는데, 이는 통행발생패턴의 비선형성에 기인하는 것으로, 인간의

〈표 3〉 결정계수(R^2) 변화추이

| 모형 | 입력 변수 | | | | |
|-----------|-------|------|------|------|------|
| | 11 | 12 | 13 | 14 | 15 |
| 선형 회귀 | 0.50 | 0.47 | 0.40 | 0.38 | 0.30 |
| 지수 회귀 | 0.63 | 0.64 | 0.63 | 0.62 | 0.58 |
| 비선형 회귀 모형 | 0.59 | 0.58 | 0.59 | 0.59 | 0.60 |
| 신경망 모형 | 0.84 | 0.78 | 0.85 | 0.89 | 0.85 |



〈그림 5〉 결정계수(R²) 변화추이

의사결정 과정을 모사한 신경망 모형의 우수성에 기인한 것으로 판단된다. 비선형 회귀모형의 경우 통행 발생패턴의 비선형성을 추정하기 보다는 모형에 내재된 비선형성에 조사된 자료를 역으로 적합시키는 형태이므로, 추가정보에 따라 결정계수가 개선되지 않는 결과를 보여주고 있다.

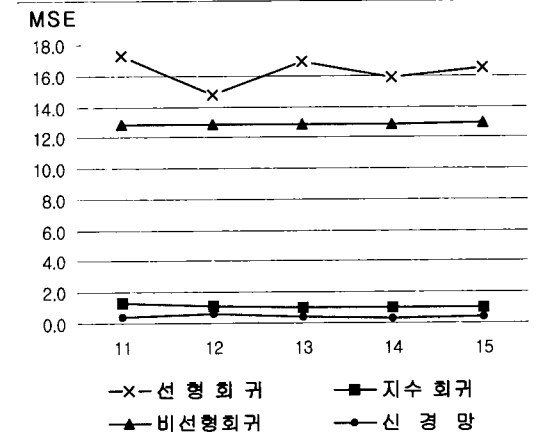
3) 오차 및 분산도

〈표 4〉과 〈그림 6〉은 평균제곱오차(MSE)의 변화추이를 나타낸 것이다. MSE(Mean Squared Error)의 변화추이를 보면 회귀모형 중에서는 지수모형이 가장 우수하게 나타나고 있다. 비선형 회귀모형이나 선형 회귀모형과 비교할 때 비교적 안정적이면서도 낮은 오차의 변화도를 보여주기 때문이다. 이는 모형의 추정과정에서 로그변환을 통한 비율의 조정에 기인한 것으로 모형에 내재된 장점이라기 보다는 추정과정에서 얻은 부산물의 성격이 좋다. 이에 비해 신경망 모형은 여타 회귀모형과 비교할 수 없을 정도의 우수한 MSE값의 변화도를 나타내고 있다. 모든 추가정보에 대해 오차 0.5이하의 우수한 검정력을 보여주고 있다. 반면 선형 회귀모형은 높은 MSE값을 나타내, 모형의 오차가 큰 것으로 나타났다.

〈표 5〉과 〈그림 7〉은 평균 상대 분산도(ARV)의 변화추이를 나타낸 것이다. ARV(Average Relative Variance)의 경우도 MSE와 비슷한 결과를 보여주고 있다. 〈표 5〉에서 볼 수 있듯이 신경망 모형의 상대적 분산도가 낮아 교사치가 목표값 주변에 고르게 분포되어 있음을 보여주고 있다. MSE의 경우처럼

〈표 4〉 MSE 변화추이

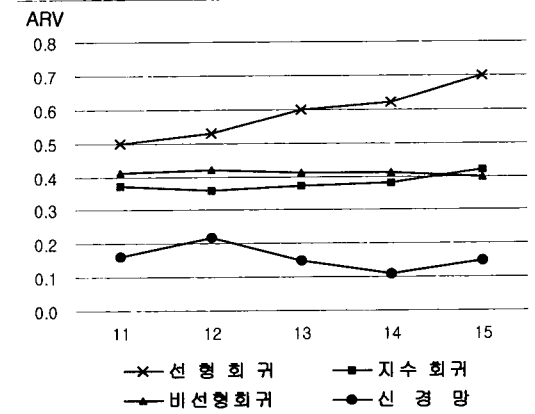
| 모형 | 입력 변수 | | | | |
|----------|-------------|-------------|-------------|-------------|-------------|
| | 11 | 12 | 13 | 14 | 15 |
| 선형 회귀 | 17.30 | 14.80 | 16.90 | 15.90 | 16.50 |
| 지수 회귀 | 1.31 | 1.16 | 1.05 | 1.03 | 1.04 |
| 비선형 회귀모형 | 12.84 | 12.85 | 12.86 | 12.86 | 12.90 |
| 신경망 모형 | 0.43 | 0.58 | 0.44 | 0.35 | 0.38 |



〈그림 6〉 평균제곱오차(MSE) 변화추이

〈표 5〉 ARV 변화추이

| 모형 | 입력 변수 | | | | |
|----------|-------------|-------------|-------------|-------------|-------------|
| | 11 | 12 | 13 | 14 | 15 |
| 선형 회귀 | 0.50 | 0.53 | 0.60 | 0.62 | 0.70 |
| 지수 회귀 | 0.37 | 0.36 | 0.37 | 0.38 | 0.42 |
| 비선형 회귀모형 | 0.41 | 0.42 | 0.41 | 0.41 | 0.40 |
| 신경망 모형 | 0.16 | 0.22 | 0.15 | 0.11 | 0.15 |



〈그림 7〉 평균상대분산도(ARV) 변화추이

지수모형의 상대적 분산도가 비선형 회귀모형에 비해 우수한 값을 나타내고 있다. 이 역시 로그변환을 통한 비율의 조정효과로 풀이된다.

3. 신경망 모형의 교사치²⁾ 및 추정치³⁾ 분석

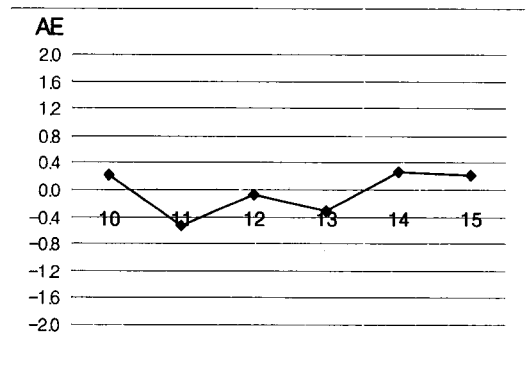
1) 신경망 모형의 교사치 분석

신경망 모형의 교사치를 검토하기 위해 평균오차(AE: Average Error)와 적합도를 분석한다. 평균오차는 추가정보에 따른 모형의 안정도를 보여주는 척도가 될 수 있다. 정보가 추가될 때마다 오차가 심하게 변동된다면 우수한 모형이 될 수 없고, 고른 오차 분포를 나타낸다면 우수한 모형이라 할 수 있다.

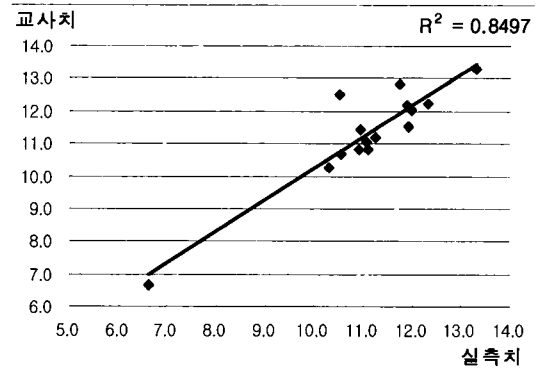
$$AE = \frac{\sum_{i=1}^N (\hat{Y}_i - Y_i)}{N} \quad (14)$$

〈그림 8〉은 신경망 모형의 평균오차(AE) 변화추이를 나타낸 것이다. 신경망 모형은 평균오차가 ±0.5 이내의 아주 우수한 수치를 보여주고 있다. 이는 교사치에 그대로 반영되어 실제 통행발생량과 거의 선형에 가까운 관계를 나타내고 있다.

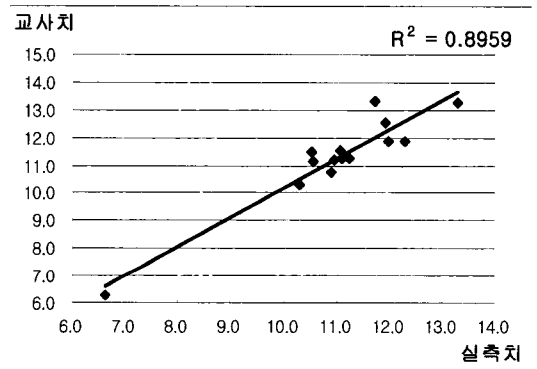
〈그림 9〉과 〈그림 10〉은 신경망 모형으로 추정된 자료 중 15개존과 14개존이 포함된 경우의 실제 통행발생량과 모형내에서 추정된 통행발생량(교사치)과의 관계를 나타낸 것이다. 두 경우 모두 완벽한 추정(R²=1.0)에 근사한 결과를 나타내었다.



〈그림 8〉 평균오차(AE) 변화추이



〈그림 9〉 15개 존 자료



〈그림 10〉 14개 존 자료

2) 신경망 모형의 추정치 분석

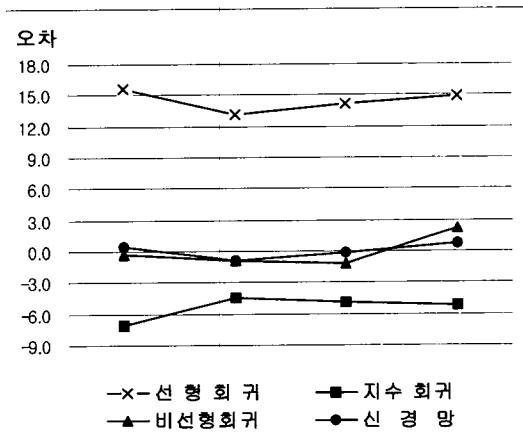
모형간 추정치의 비교 검토는 학습에 이용되지 않은 자료를 이용한다. 보통의 신경망 연구에서는 사전에 자료의 일부분을 검정자료로 할당하고, 이를 학습된 모형(네트워크)에 적용해 결과를 분석한다. 본 연구에서는 전체 관측치가 15개존이라는 제한된 조건으로 인해 자료를 사전에 할당하는 통상적 방법을 적용하지 않았다. 연구에 적용된 방법은 11개존을 대상으로 한 모형에서, 선택되지 않은 4개존을 추정치 검정자료로 이용했다. 또한 12개 이상 존자료의 경우 검정자료의 부족(3개이하)으로 추정치 분석을 생략한다.

〈표 6〉은 모형간 추정치의 적합성 차이를 보여주고 있다. 표에 제시된 자료는 실측치와 추정치의 오차를 나타낸 것이다. 선형 회귀모형의 경우 13.07~15.65 사이의 값을 나타내고 있으며, 지수모형의 경우 -7.15~-4.44 사이의 값을 나타내고 있다. 비선형

2) 교사치는 학습에 사용되었던 출력단위의 값으로 정의한다.
3) 추정치는 학습된 신경망 모형을 이용한 검정자료의 출력치로 정의한다.

〈표 6〉 추정치 분석결과

| 모형 | 검정준 | | | |
|----------|-------------|--------------|--------------|-------------|
| | 1 | 5 | 11 | 14 |
| 선형 회귀 | 15.65 | 13.07 | 14.17 | 14.94 |
| 지수 회귀 | -7.15 | -4.44 | -4.81 | -5.15 |
| 비선형 회귀모형 | -0.33 | -0.95 | -1.13 | 2.24 |
| 신경망 모형 | 0.50 | -0.89 | -0.16 | 0.80 |



〈그림 11〉 추정치의 오차

모형은 -1.13~2.24 사이를 나타내며, 신경망 모형은 -0.89~0.80 사이의 값을 나타내었다. 결과에서 보듯이 선형 회귀모형은 신경망 모형에 비해 오차가 14~15배에 달하며, 지수모형은 7~8배에 달하고 있다. 비선형 모형은 신경망 모형에 가장 근접한 오차 범위(2배)를 나타내었다. 신경망 모형은 교사치 뿐만 아니라 추정치에서도 우수한 결과를 나타낸다는 것이 실증적으로 확인된 것이다.

V. 결론

본 연구는 『제1차 전국물류현황 조사』 자료를 이용하여 기존의 대표적인 통행발생모형인 회귀모형과 신경망 이론에 의한 통행발생모형과의 결과를 비교·분석하여 통행발생모형에 대한 새로운 방법을 제시하고자 하는 것이었다. 이를 위해 회귀분석 일색의 기존 연구의 한계를 밝히고, 신경망 모형의 우수성을 입증하고자 시나리오별 모형의 안정성 분석을 통한 각종 통계적 검정을 시도하였다. 연구의 결과는 다음과 같다.

첫째, 회귀모형은 모수치 추정에서 작게는 2~3배의 편차를 보여 불안정한 모습을 보였다. 특히 선형 회귀모형은 큰 편차에 비상식적인 부호를 나타내어 불합리한 패턴을 나타내었다. 지수모형의 경우는 로그변환을 통한 비율조정의 효과로 선형 회귀모형에 비해 비교적 안정된 양상을 나타내었다. 비선형 회귀모형은 회귀모형 중 가장 안정된 모수치 편차와 적절한 변수 부호를 나타내었다.

둘째, 모형의 전반적 검정력을 보여주는 결정계수에서는 신경망 모형의 우수성이 두드러졌다. 평균 0.8을 상회하는 우수한 검정력에 안정된 패턴을 보여 변량에 대한 설명력이 가장 높았다. 반면, 회귀모형은 비선형 모형만이 비교적 안정된 패턴을 보여주었을 뿐, 선형회귀모형과 지수모형 모두 채택하기 어려운 통계치를 나타내었다. 이는 통행발생 패턴의 복잡함, 즉 인간의 의사결정에 기반한 비선형성이 모형간 검정력의 차이로 귀결되었음을 의미한다.

셋째, 모형의 정밀함은 오차와 분산도로 측정하였다. MSE의 경우 역전파 모형은 선형회귀모형의 1/40, 지수모형의 1/3, 비선형회귀모형의 1/20수준으로 아주 우수한 결과를 보여주었다. ARV 역시 회귀모형의 1/2수준의 양호한 통계치를 나타내었다. 이는 예측값이 목표값 주변에 고르게 분포되어 있음을 의미하는 것으로, 본 연구에 의하면 85% 이상의 설명력(〈그림 9〉, 〈그림 10〉)이 있는 것으로 드러났다. 이는 비선형 사상(Nonlinear mapping)이 가능한 역전파 모형의 우수성이 실증적으로 입증된 것이라 하겠다.

넷째, 모형의 장래 예측력을 판단할 수 있는 추정치 분석에서도 신경망 모형의 우수성이 실증적으로 입증되었다. 검정 자료를 이용한 오차분석에서 회귀모형은 신경망 모형의 오차범위를 2~7배 이상 상회하는 양상을 나타내었다. 반면, 신경망 모형은 오차범위 ±1.0 내의 안정된 추정치를 나타내었다.

본 연구의 결과, 역전파 신경망 모형은 기존 회귀모형의 검정력과 안정성을 훨씬 상회하는 결과를 보여 주었으며, 이는 향후 통행발생모형으로 역전파 신경망 모형의 적극적 활용 가능성을 의미하는 것으로 해석할 수 있다. 특히 더욱 복잡해진 교통현상과 수집자료의 다양성을 고려할 때 교통분야에서의 신경망 모형의 적용은 더욱 확대될 것으로 전망된다.

참고문헌

1. Box and Cox(1964), "An Analysis of Transformations", Journal of the Royal Statistical Society: Series B, pp.211~243.
2. Chatterjee S. and B. Price(1991), "Regression Analysis by Example", 2nd, Wiley.
3. Gaudry, M. J. I. and M. G. Degenais(1979), "The Dogit Model", Transportation Research, 13, B.2, pp.105~112.
4. J. Barcelo and J. Casas(1999), "The Use of Neural Networks For Short-Term Prediction Of Traffic Demand", Transportation and Traffic Theory, Pergamon.
5. Laurene Fausett(1984), "Fundamentals of Neural Network", Prentice Hall.
6. Mark Dougherty(1995), "A Review of Neural Networks Applied to Transport", Transportation Research-C, Vol. 3, No. 4, pp.247~260.
7. Mandel B. et al(1994), "Linear or Nonlinear Utility Functions in Logit Models? The Impacts on German High-Speed Rail Demand Forecasts", Transportation Research-B, Vol. 28B, No. 2. pp.91~101.
8. McCulloch, W. S. and Pitts, W.(1943), "A logical calculus of the ideas immanent nervous activity", Bulletin of Mathematical Biophysics, 5th, pp.115~133.
9. Oum T. H.(1989), "Alternative Demand Models and Their Elasticity Estimates", JTEP, May, 1989, pp.163~187.
10. Picard G. and M. Gaudry(1998), "Exploration of a Box-Cox Logit Model of Inter-city Freight Mode Choice", Transportation Research-E, Vol. 34, No. 1. pp.1~12.
11. Reprinted in J. A. Anderson and E. Rosenfeld (Eds)(1988), "Neurocomputing: Foundations of Research", Cambridge, Massachusetts: MIT Press.
12. Rumelhart, D. E., McClelland, J. L. and the PDP Research Group (Eds)(1986a), "Parallel Distributed Processing: Explorations in the Microstructure of Cognition", Vol. 1 & 2, Cambridge, Massachusetts: MIT Press.
13. Velihimanan et al(1998), Neural Networks in Transport Application, Ashgate.
14. Yang H. et al(1993), "Exploration of Driver Route Choice with Advanced Traveler Information Using Neural Network Concepts", Transportation Research Board 72nd Annual Meeting January 10~14.