

■ 論 文 ■

단기조사 교통량을 이용한 AADT 추정연구

A study on the estimation of AADT by short-term traffic volume survey

이 승 재

(서울시립대학교 교통공학과 부교수)

백 남 철

(한국건설기술연구원 연구원)

권 희 정

(서울시정개발연구원 연구원)

목 차

I. 서론

1. 연구의 배경 및 목적
2. 연구의 범위

II. 상시조사지점의 변동요인분석

III. 군집분석

1. 요인의 결정
2. 그룹수 결정

IV. 그룹할당

1. 판별분석을 이용한 그룹할당의 단계
2. 신경망 이론을 이용한 그룹할당의 단계
3. 그룹할당의 결과

V. 결론

참고문헌

Key Words : AADT 추정, 군집분석, 그룹할당, 판별분석, 신경망

요 약

도로의 설계 및 운영 등에 필요한 연평균 일 교통량은 365일 조사에 의한 것이 아닌 단기간 조사된 교통량을 사용하는 것으로써 이를 추정하려는 연구는 이전부터 있어왔다. 본 연구에서는 기존 연구를 바탕으로 이 AADT 추정의 방법을 개선시키고자 하였다.

먼저 그룹간의 차이를 뚜렷이 보여줄 수 있는 변수를 찾기 위해 그룹의 수를 변화시켜가며 각 그룹의 시간변동요인들(전체, 주중, 토요일, 일요일, 주중-토요일, 주중-일요일)의 값을 살펴보아 그 차이가 가장 뚜렷한 변동요인을 주중-일요일의 시간변동 요인으로 선정하였다.

그 다음 월변동요인만을 사용하여 상시조사지점을 clustering하였다. 그룹간의 시간변동요인의 차이를 가장 크게 하는 것을 원칙으로 군집분석을 한 결과 10개의 그룹으로 묶을 수 있었다. 선정된 주중-일요일의 시간변동요인을 사용하여 판별분석과 신경망을 통한 그룹할당을 했다. 신경망의 적중률이 판별분석의 경우보다 훨씬 좋았고, RMSE, U-test 결과도 더 좋았다.

결과를 전체적으로 살펴보면, 본 연구에서 사용한 방법(월변동요인만을 사용하여 군집분석한 후, 각 그룹에서 월별로 요일변동요인을 구해 적용한 AADT 추정)의 결과가 이전 연구인 월변동과 요일변동을 이용한 AADT 추정의 결과보다 훨씬 좋았다. 그리고 그룹할당의 변수를 주중-일요일의 시간변동요인으로 달리하였을 때, 신경망의 경우 그룹할당의 적중률이 더 높아지는 것을 볼 수 있었다.

1. 서론

1. 연구의 배경 및 목적

연평균 일 교통량(Annual Average Daily Traffic: AADT)은 도로의 계획, 설계 및 유지관리를 위한 기초자료로 사용되는 값이다. AADT가 과소 추정되면 시공되는 차로 수가 모자라며, 과대 추정될 경우에는 도로 시설의 일부가 낭비되는 결과를 초래한다.

이러한 AADT를 산출하기 위해서는 정기적인 도로교통량 조사가 필요한데, 이를 위해 현재 우리나라에서는 상시조사(permanent count)와 전역조사(temporary 또는 coverage count)로 구분하여 교통량 조사를 시행하고 있다.

상시조사지점에서는 1년 365일간 교통량이 조사되므로, 연간 수집된 교통량을 365로 나누면 AADT가 산출된다. 이에 반해, 전역조사지점에서 이루어지는 단기조사는 미국 일본 등과 같이 2일 동안 수행하고 있으며, [단, 미국 Traffic Monitoring Guide(1985, 1992, 1995, 2000)에서는 48시간조사를 원칙으로 하면서도, 주별 사회경제적 도로인프라 특성에 따라 그리고 도로구간특성에 따라 조사기간을 다르게 적용하고 있다.] 따라서 AADT가 추정되어야 한다.

또한, 조사 요일을 보면 불규칙적인 교통량을 보이는 일요일은 단기조사기간에 포함되지 않는 것이 일반적이는데, AADT에 근접하여 일상적인 교통량을 나타낸다고 분석된 화, 수, 목요일이 교통량 조사일로 선호되고있는 실정이다.

본 논문에서는 단기조사는 평일에 기본적으로 이루어진다고 보면서, 특성이 뚜렷한 주말 교통량도 함께 분석하여 패턴그룹을 어떤 방법으로 판단하는 것이 타당한지에 대해 살펴보았다.

AADT를 추정하는 방법 중에는 조사된 몇 일의 교통량을 단순 평균하는 방법과, 가장 가까운 상시조사지점의 AADT를 그대로 쓰는 방법, 그리고 유사한 교통패턴을 보이는 지점의 조정계수를 가져다가 일평균교통량에 보정해 주는 방법이 있다. 여기서는 보정을 이용한 방법론을 연구해보고자 한다.

즉, 기존의 월변동과 요일변동요인을 사용한 AADT 추정의 단점을 보완하여 그룹할당이나 교통량 보정에 대한 효과를 높이고자 함이다. 그렇게 함으로써 궁극적으로는 전역조사지점의 AADT 추정치의 신뢰성을

높이고자 하는 것이다.

2. 연구의 범위

논문에서 사용한 자료는 건설기술연구원에서 상시조사지점으로 선정한 171개 지점에 대해 1998년, 1999년에 걸쳐 365일 조사된 자료이다. 따라서 이 지점은 연평균일교통량 실측치와 월변동요인, 시간변동요인을 모두 갖게된다. 이러한 지점의 자료를 이용해서 군집분석을 하고 그룹으로 나누었으며, 그 결과 각 지점은 소속그룹에 대한 정보도 또한 갖게된다.

두 번째 단계인 그룹할당의 단계에서도 같은 171개의 지점을 사용하였는데, 이렇게 실측 AADT자료와 소속그룹자료를 갖는 지점의 교통량을 이용함으로써 본 연구에서 사용한 AADT추정의 정확성을 검증할 수 있게 되는 것이다.

본 연구의 내용은 교통량 패턴분석, 상시조사 지점의 clustering, 해당 그룹으로의 전역조사 지점의 할당과 AADT 추정의 과정으로 이뤄지며 각 과정의 분석 내용은 다음과 같다.

1) 상시조사 지점의 clustering - 패턴그룹 설정

먼저 각 조사지점의 교통패턴을 분석한다. 이때 분석되는 사항은 월변동, 시간변동요인이다.

조사된 변동요인들 중 월변동요인을 이용해 군집분석을 행하고, 각 그룹의 시간변동요인 값들을 살펴보고 그룹간 차이를 가장 크게 하는 그룹수로 결정한다.

2) 패턴그룹 할당 방법의 평가

상시조사지점을 전역조사 지점이라 가정하고 미리 설정해 놓은 교통량 패턴그룹으로 할당한다. 이때 시간변동요인 24개를 이용하는 판별분석, 신경망 방법을 사용한다.

3) AADT 추정과 모형의 평가

전역조사지점이라 가정된 상시조사지점의 AADT를 추정한다. 이때 사용되는 값은 월변동요인, 요일변동요인 그리고 알고자 하는 지점의 실측 ADT이다.

II. 상시조사지점의 변동요인분석

선별된 유효데이터를 대상으로 지점별 월변동, 요일변동, 시간변동 패턴을 구한다.

i, j : 각각, 월, 요일
 FM_{ni} : n 지점 i 월의 월변동요인
 FD_{nj} : n 지점 j 요일의 요일변동요인
 FH_{nl} : n 지점 l 시간대의 시간변동요인

$$FM_{ni} = \frac{MADT_{ni}}{AADT_n} \quad (1)$$

$$FD_{nj} = \frac{\sum_{k=1}^{n(j)} DT_{nkj}}{\text{일년중 } j\text{요일의 일수}(n(j))} \quad (2)$$

(DT_{nkj} : n 지점 k 번째 j 요일의 교통량)

$$FH_{nl} = \frac{\sum_{\forall j(=\text{월일})} HT_{n \cdot jl}}{\text{연중 평일의 일수}} \quad (3)$$

(HT_{nl} : n 지점 j 번째날 l 시간대의 시간교통량)

여기서,

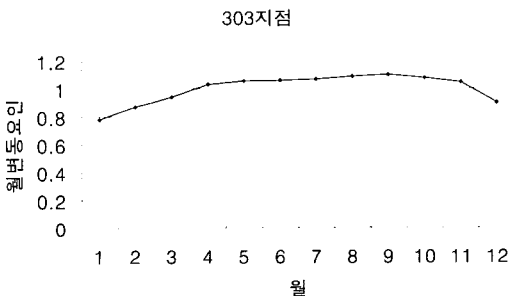
$$MADT_{ni} = \frac{\sum_{j=1}^{n(i)} DT_{nij}}{n(i)} \quad (4)$$

($n(i)$: i 월의일수)

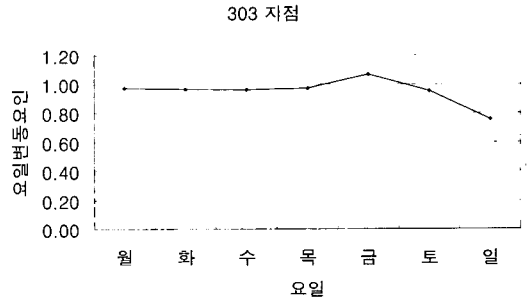
$$AADT_n = \frac{\sum_{j=1}^{365} DT_{nj}}{365} \quad (5)$$

(DT_{nj} : n 지점 j 번째날의 일교통량)

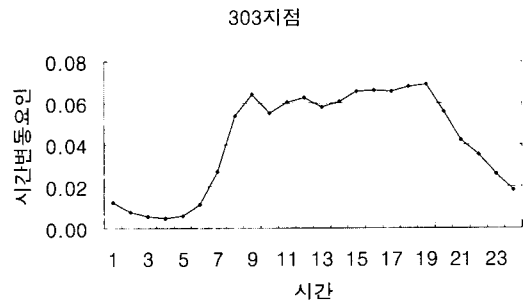
실례로 303지점(경기도 광주)에 대한 변동요인 값들을 구해보았다. 월변동요인은 <그림 1>, 요일변동요인은 <그림 2>, 시간변동요인은 <그림 3>에 나타나 있다.



<그림 1> 303지점의 월변동요인



<그림 2> 303지점의 요일변동요인



<그림 3> 303지점 시간변동요인

III. 군집분석

1. 이론과 적용

군집분석(cluster analysis)은 모집단 또는 범주에 대한 사전정보가 없는 경우에 관측치(observation)들 사이의 유사성-거리 혹은 상관관계-을 이용하여 전체를 몇 개의 군집(cluster)으로 나누는 통계적 분석기법이다. 결과적으로 그룹 내에서 동일한 특성을 나타내고, 그룹간에는 서로 상이한 특성을 나타내도록 모집단을 여러 그룹으로 나누는 방법이라 할 수 있다.

군집분석은 판별분석과 다르다. 판별분석은 분류하기 전에 미리 그룹수를 결정한 후 새로운 관찰대상을 이미 정해진 집단들 중의 하나에 할당하는 것을 목적으로 하는 반면, 군집분석은 그룹의 수를 미리 정하지 않고 전체 대상들에 대한 유사성이나 거리에 의거하여 동질적인 집단으로 분류하는 것을 목적으로 한다.

군집분석에서 결정되어야 할 문제는 크게 유사성 척도, 군집화 방법, 그룹수의 결정으로 나누어진다.

단기조사 교통량으로 AADT를 추정했던 한국건설기술연구원(2000)의 연구보고서에서는 월변동, 요일변동으로 각각 그룹을 구분한 후 이를 조합하는 방식으로 그룹을 구분하였다.

그런 후 시간변동요인을 사용하여 그룹할당을 하는데, 여기에 문제가 있다. 월변동, 요일변동을 이용해 나누어놓은 그룹간의 시간변동요인 값들이 크게 차이를 보이지 않았던 것이다. 그룹할당의 기준이 되는 값들에서 그룹간의 뚜렷한 차이가 나지 않으면, 그룹할당시에 오차가 발생할 확률이 커지게 된다.

본 연구에서는 월변동요인만으로 그룹을 구분하고자 하였다. 그룹을 나누는 방식은 군집분석의 여러 방법 중 일반적으로 가장 많이 사용되고있는 Ward법을 사용하였다. Ward법의 적정성은 미국 Traffic Monitoring Guide(1985, 1992, 1995, 2000)에서 연구결과로 확인되었다.

군집분석에서는 그룹수를 대략 2개~15개 정도로 설정하여 각 경우마다 그룹별 시간변동요인을 구하였다. 이렇게 구해진 그룹별 시간변동요인이 가장 큰 차이를 보이는 때를 그룹설정의 최적의 상태로 두었고 이 그룹수를 기준으로 상시조사지점을 구분하였다.

군집분석의 그룹수 결정에는 최종적으로 RMSE와 U-test 값을 사용하게 되는데 일반적으로 이 값들은 그룹의 수가 증가할수록 낮아지는 추세를 보이게 된다. 낮은 값을 보일 때가 적정한 그룹수라고는 하지만 무작정 그룹의 수를 늘려서는 안되겠기에, 사전에 그룹의 범위를 좁히는 작업을 하게 된다. 이때 사용되는 값이 R^2 와 군내분산 S_{yy} 그리고 Pseudo F값이다.

한편, 시간변동요인도 그 특성에 따라 여러 가지로 구분할 수가 있다. 즉, 전체, 주중, 토요일, 일요일, 주중-토요일, 주중-일요일로 구분가능한데, 이들 중에서 그룹을 변화시켰을 때 그룹별로 가장 뚜렷한 차이를 보이는 값으로 선정하였다.

사용할 시간변동요인이 결정된 후 월변동요인을 사용하여 그룹을 묶었고, 각 단계에서 시간변동요인에 대한 통계치를 살펴보았다.

2. 요인의 결정

시간변동요인은 여러 가지 종류로 나눌 수 있다. 즉 주중시간변동, 토요일시간변동, 일요일시간변동,

주중-토요일 시간변동, 주중-일요일 시간변동인데, 이는 시간변동요인의 분포를 살펴보아 구분한 것이다.

월~금까지의 주중 시간변동요인은 유사한 분포를 나타내었고, 토요일 및 일요일의 시간변동요인의 분포가 각각 차이를 보였으며, 주중의 시간변동요인과의 차이도 보였다. 그 결과에 의해 크게 세 가지로 구분하였고 가장 큰 차이를 보이는 주중시간변동요인을 나머지 두 시간변동요인과 각각 결합하여 조합을 만들었다.

이들을 이용하여 어느 것을 사용할 때 가장 그룹별 특성이 뚜렷한가를 살펴보는 것이 필요하다.

방법을 살펴보면,

- ① 월변동요인을 이용하여 군집분석한다.
- ② 그룹수는 2개에서 15개까지 변화시킨다.
- ③ 월변동요인으로 나누어진 각 그룹의 시간변동요인에 대해 R^2 와 군내분산 S_{yy} 그리고 Pseudo F값을 구한다. R^2 , Pseudo F는 높고, S_{yy} 값은 작은 변동요인을 선택한다.

$$S_{yy} = \sum_i (HT_{ij} - \overline{HT}_i)^2 \quad (6)$$

$$R^2 = \frac{\sum_i \sum_j (HT_{ij} - \overline{HT})^2 - \sum_i \sum_j (HT_{ij} - \overline{HT}_i)^2}{\sum_j n_j (\overline{HT}_j - \overline{HT}) + \sum_i \sum_j (HT_{ij} - \overline{HT}_i)^2} \quad (7)$$

여기서,

i, j : 각각, 관측지점, 그룹 번호

n_j : 그룹 j의 시간변동요인 수

HT_{ij} : 그룹 j에 속한 지점 i의 시간변동요인 관측치

\overline{HT}_i : 지점 i가 속한 그룹의 시간변동요인 평균치

\overline{HT} : 대상이 된 모든 시간변동요인의 평균치

각 시간변동요인에 대한 값을 보면 일요일과 주중-일요일의 시간변동이 Pseudo F와 R^2 에서 가장 높았고, 그래프로 제시되지는 않았지만 주중-일요일의 시간변동이 S_{yy} 값에서 가장 작았다. 세 가지 결과로 볼 때 주중-일요일과 일요일의 시간변동요인 모두 적합해 보이기 는 하지만, 기본적으로 단기조사는 주중에 이뤄져야하므로 두 경우 중 주중-일요일의 시간변동

요인을 선택하였다.

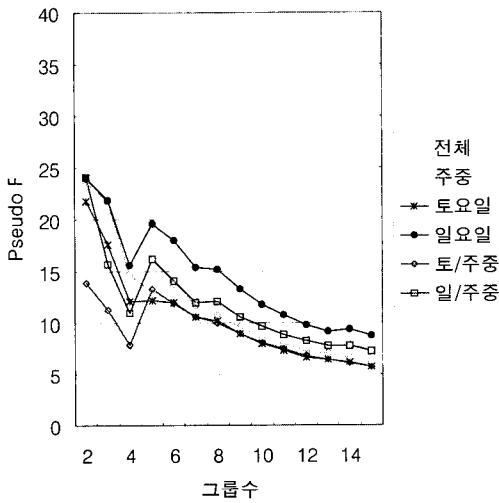
1998, 1999년도의 자료를 이용하여 월그룹 2~15일 때의 시간변동요인에 대한 값들의 분포를 살펴 보았다.

1998년의 pseudo F와 R^2 의 경우(〈그림 4〉, 〈그림 5〉)는 전체적으로 일요일, 전체, 일요일-주중의 시간변동

순으로 높게 나타났다.

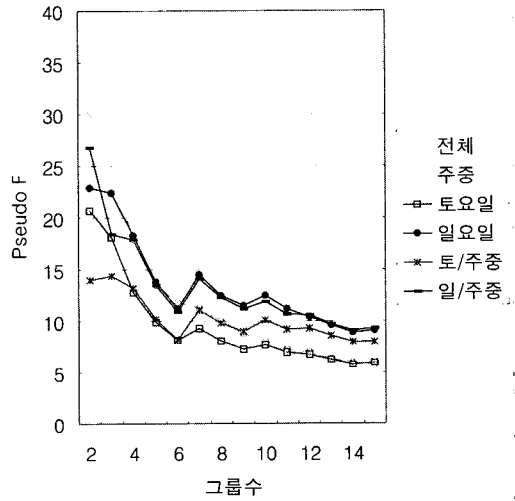
또, 1999년도의 pseudo F와 R^2 (〈그림 6〉, 〈그림 7〉) 값을 보면, 그룹수 6개까지는 전체, 주중, 일요일 시간변동의 순으로 높게 나타났지만, 6개 이후에는 일요일, 일요일-주중 시간변동의 순으로 높은 값을 보였다.

1998년 자료 - Pseudo F



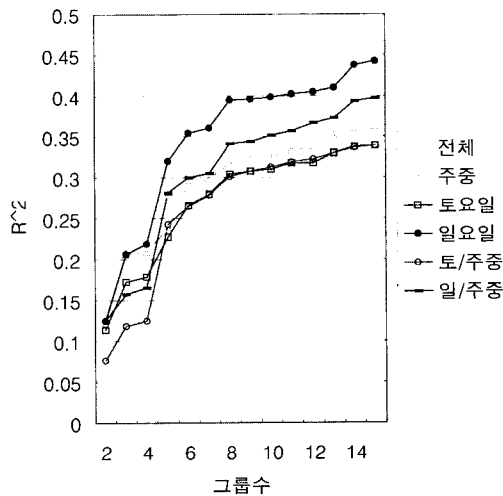
〈그림 4〉 그룹수 변화에 따른 Pseudo F

1999년 자료의 Pseudo F



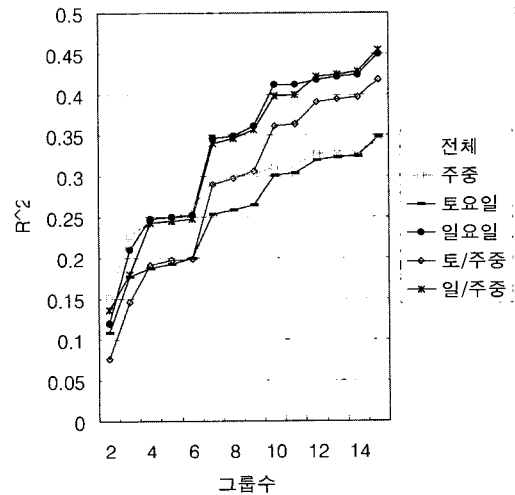
〈그림 6〉 그룹수 변화에 따른 Pseudo F

1998년 자료의 R^2



〈그림 5〉 그룹수 변화에 따른 R^2

1999년 자료의 R^2



〈그림 7〉 그룹수 변화에 따른 R^2

3. 그룹수 결정

적정 그룹수는 선정된 주중-일요일의 시간변동요인에 대한 통계치를 보고 결정을 하게된다. 그룹을 나누는 원리는 그룹별 주중-일요일 시간변동요인의 차를 가장 뚜렷하게 하는 그룹수를 찾는 것이다. R^2 는 크게, 군내분산 Syy 는 작게, Pseudo F값은 국부적 극대값을 나타내는 그룹수를 최적의 그룹수로 정한다.

Syy 값은 그룹수가 늘어날수록 전체적으로 감소하는 형태를 보여 그룹수 결정에는 크게 영향을 미치지 않아 보이므로 앞에서 보여진 <그림 4, 5, 6, 7>을 보고 개략적인 그룹수를 정한다.

R^2 는 1998년의 경우 대체적으로 6개 이상부터 경사가 완만해지고, 1999년의 경우 8개 이상부터 경사가 완만해 진다. Pseudo F는 1998년의 경우 그룹수 5개에서 국부적 극대값을 보이고, 1999년의 경우 7개, 10개에서 국부적 극대값을 보인다. 그 결과를 종합하여보면 그룹수 5개에서 10개 사이에서 최적 그룹수를 결정할 수 있다.

최종적으로 그룹수를 결정하기 위해 각 경우의 AADT를 추정한 후, RMSE나 U-test를 구해 그 값을 가장 작게 하는 그룹수로 결정하기로 한다.

▷ RMSE

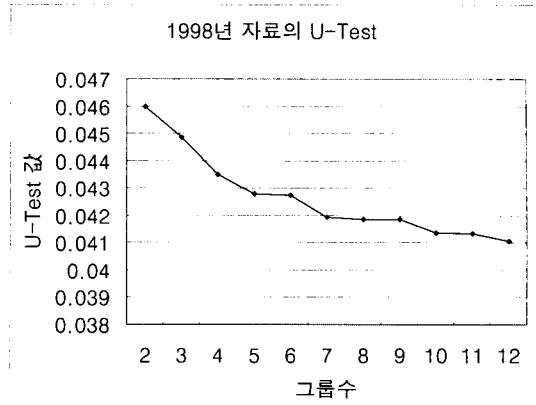
$$RMSE_n = \sqrt{\frac{\sum_{i=1}^N (\widehat{AADT}_{mi} - AADT_n)^2}{N}} \quad (8)$$

▷ U-TEST

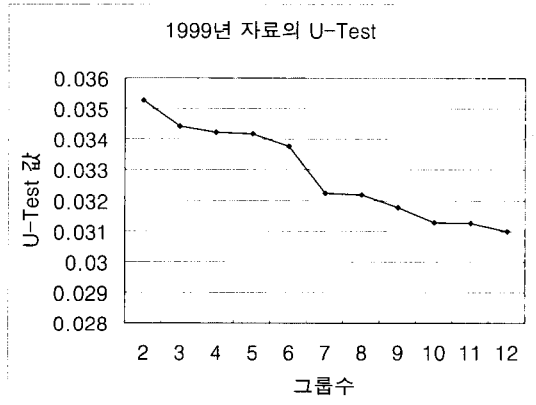
$$U_n = \frac{\sqrt{\frac{\sum_{i=1}^N (\widehat{AADT}_{mi} - AADT_n)^2}{N}}}{\sqrt{\frac{\sum_{i=1}^N (\widehat{AADT}_{mi})^2}{N}} + \sqrt{\frac{\sum_{i=1}^N (AADT_n)^2}{N}}} \quad (9)$$

두 분석 결과치가 절대적인 크기에서는 다르나 전체적인 형태는 비슷하므로 U-test 결과만 <그림 8>과 <그림 9>의 그래프로 도출하였다.

앞에서 개략적으로 정한 5개~10개 사이의 군집수를 대상으로 U-test 값의 변화를 보면, 10개일 때가 가장 적절한 것으로 나타난다. 따라서 그룹수는 10개로 결정짓는다.



<그림 8> 그룹수 변화에 따른 U-test(1998년자료)



<그림 9> 그룹수 변화에 따른 U-test(1999년자료)

IV. 그룹할당

주중-일요일의 시간변동요인을 이용하여 월변동요인과 시간변동요인을 이용해서 미리 설정해 놓은 상시조사지점들의 그룹으로 그룹할당을 한다. 그룹할당은 신경망과 판별분석을 이용한 방법 각각을 사용하였으며 후에 적응률을 비교하였다. 그리고 이 결과와 관련하여 AADT를 추정한 후, RMSE와 U-test를 구하여 평가하였다.

1. 판별분석을 이용한 그룹 할당의 단계

1) 판별분석

Fisher에 의해 개발된 판별분석은 계량적으로 측정된(등간척도나 비율척도) 독립변수들을 이용하여 명목척도로 측정된 종속변수 즉, 집단의 분류를 하는

방법이다. 이것은 분류되어 있는 집단간의 차이를 설명해 줄 수 있는 독립변수들을 찾아내고 이들의 선형 결합으로 판별식을 만들어 내어 이 식에 분류하고자 하는 각 대상들의 특성을 대입해서 각 대상들이 속하는 집단을 찾아내는 것이다.

이러한 판별분석은 개별적인 변수만을 이용하여 집단을 분류할 때보다 영향을 미치는 여러 개의 변수들을 동시에 고려하여 집단간의 차이를 나타내주는 장점을 가진다. 따라서 보다 의미 있는 결과를 나타내는 경우가 많다.

집단들간의 차이를 설명해 주기 위한 기준으로서 판별분석에서는 집단간 분산/집단내 분산의 비율을 사용하게 되는데, 이를 극대화시키는 독립변수들의 선형결합을 판별식이라 한다.

일반적인 판별함수의 형식은 식(10)과 같다.

$$Z = W_1 X_1 + W_2 X_2 + W_3 X_3 + \dots + W_k X_k \quad (10)$$

Z : 판별점수

W_i : i 번째 독립변수의 가중치

X_i : i 번째의 독립변수

k : 독립변수의 수

판별분석을 적절히 활용하기 위해서는 첫째, 각각의 모집단이 다변량 정규분포(multivariate normality)를 이루어야 한다는 것과 둘째, 모집단의 분포 및 공분산 구조(dispersion and covariance structure)가 같아야 한다는 가정이 지켜져야 한다. 또한 분류(classification)의 정확성을 결정할 때도 모든 대상에 대해 잘못 분류함으로써 생기는 비용이 동일하다는 것과 각 집단의 분포와 공분산구조도 알아야 한다는 가정이 지켜져야 한다.

2) 판별분석의 단계

Step 1) 판별함수의 작성

$$Z_{nk} = a_{k0} + \sum_{p=1}^{24} a_{kp} \cdot x_{np} \quad (11)$$

Z_{nk} : k그룹에 속한 n지점의 판별함수 값

x_{np} : n지점의 p시간의 시간변동요인

판별 상수

$$a_{k0} = - \frac{\sum_{p_1=1}^{24} \sum_{p_2=1}^{24} X_{kp_1} \cdot D(p_1, p_2)^{-1} \cdot X_{kp_2}}{2} \quad (12)$$

판별 계수

$$a_{kp} = \sum_{q=1}^{24} D^{-1}(p, q) \cdot X_{kq} \quad , p=1, 2, \dots, 24 \quad (13)$$

X_{kq} : k그룹의 q시간대의 시간변동의 평균

Step 2) 패턴분석 수행(지점별 주중시간변동)

x_{np} : n지점 p시간의 시간변동

Step 3) 판별점수를 계산한다.

n지점의 k군집에 대한 판별점수

$$Z_{nk} = a_{k0} + \sum_{p=1}^{24} a_{kp} \cdot x_{np} \quad (14)$$

Step 4) 판별점수가 최대인 군집에 n지점할당

$$MAX(Z_{n1}, Z_{n2}, \dots, Z_{nk})$$

Step 5) 군집 소속확률 계산

$$P_{nk} = \frac{e^{Z_{nk}}}{\sum_{k=1}^K e^{Z_{nk}}} \quad (15)$$

2. 신경망 이론을 이용한 그룹할당의 단계

1) 신경망이론

신경망은 과거 인공지능의 한계였던 학습능력의 실현을 목적으로 하는 인공지능의 연구분야로서 많은 계산단위로 이루어진 고집적 병렬망(massively parallel nets)을 이용하여 복수개의 가능성을 동시에 탐구하는 방식이다.

이러한 신경망은 패턴 인식 및 일반화, 추세예측 등에 유용하다. 이 방법은 빠르고 어떤 원칙이나 규칙을 필요로 하지 않는다는 장점 때문에 자료상의 변화에 적응시켜서 입력신호의 특징을 학습시킬 수 있다.

신경망의 기능은 다음과 같다.

① 연상(association)

신경망에 일정한 패턴이 저장될 경우 나중에 이 저장된 패턴과 유사한 패턴이 입력되면 신경망이 기억된 패턴 중 입력패턴과 가장 근사한 패턴을 찾아내는 정보처리 기능이다.

② 분류(classification)

입력패턴을 몇 개의 부류(category)로 분류하는 기능이다.

③ 일반화(generalization)

제한된 데이터를 이용하여 그 입력 데이터가 형성하고 있는 패턴을 찾아내는 기능이다.

2) 이론의 적용

본 연구에서는 신경망의 여러 모델 중에서 문자인식, 텍스트로부터 음성합성, 로봇트팔의 제어, 대출심사, 주가예측 등에 주로 사용하는 Back - Propagation 이론을 사용하도록 한다.

학습에 이용되는 유닛은 각 전역조사지점이 보유하고있는 시간변동요인들이고, 전역조사지점들을 할당할 때 쓰이는 유닛은 각각의 전역조사지점에서 조사된 단기조사교통량에서 산출한 시간변동요인들이다.

신경망이론이 수행되는 각 단계는 다음과 같다.

Step 1)

- (1) 각 유닛의 결합강도 w_{ji} , w_{kj} , 바이어스 θ_j , θ_k 의 초기치를 결정.
- (2) 학습횟수 t 의 초기치를 0, 입력벡터(데이터) p 의 초기치를 1, 학습율 η 과 관성계수 α , 학습종료 횟수 T 를 설정.

Step 2) : p 번째 입출력 데이터 (x_p , tp)를 입력.

(1) 입력층

- ① 유닛 수 : 각 지점별 시간변동요인 24개
- ② 입력 : p 지점의 시간변동요인 $x_p=(x_{p1}, x_{p2}, \dots, x_{p24})$ 로 표시

(2) 입력벡터 x_p 에 대한 교사신호 $tp=(tp_1, tp_2, \dots, tp_M)$ 는 다음과 같이 정의됨

- $tp_k=1$: 샘플 p 가 그룹 k 에 속하는 경우
- $=0$: 샘플 p 가 그룹 k 에 속하지 않는 경우

즉, 샘플 p 가 그룹 h 에 속하면 교사신호 tp 는 h 번째 요소만 1이고 나머지 요소는 모두 0임

Step 3) : 입력데이터 x_p 에 대한 각 층의 출력결과를 계산한다.

(1) 중간층

- ① 유닛 수 : 그룹수의 2배
- ② 결합강도 w_{ji} 로써 입력 unit을 결합한다.

$$netpj = \sum_{i=1}^{24} w_{ji}x_{pi} + \theta_j, \quad j=1,2,\dots,2M \quad (16)$$

(2) 출력층

- ① 유닛 수 : 그룹 수, M
- ② 결합강도 w_{kj} 로써 중간층을 결합한 값을 갖는다.

$$netpk = \sum_{j=1}^{2M} w_{kj}o_{pj} + \theta_k, \quad k=1,2,\dots,M \quad (17)$$

(3) 학습 대상이 되는 가변 매개변수

w_{ji} , w_{kj} , θ_j , θ_k (단, 바이어스인 θ_j , θ_k 는 1로 설정하는 것이 일반적)

Step 4) : 각 층의 출력결과를 이용해서 결합강도 w_{ji} , w_{kj} , 바이어스 θ_j , θ_k 를 수정한다.

(1) 각 유닛의 입출력 함수

시그모이드 함수로써,

$$\text{로지스틱 함수} : f(net) = \frac{1}{1 + e^{-net}} \quad (18)$$

(2) 입력 벡터 x_p 가 입력된 경우의 신경망의 출력과 교사신호와의 오차

$$E_p = \sum_{k=1}^M \sqrt{\frac{(t_{pk} - o_{pk})^2}{M}} \quad (19)$$

(3) BP 알고리즘 $\Rightarrow E_p$ 를 최소화하기 위한 최급 강화법

① 유닛간의 결합강도 w_{ji} 와 w_{kj} 의 수정

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji}(t+1) \quad (20)$$

$$w_{kj}(t+1) = w_{kj}(t) + \Delta w_{kj}(t+1) \quad (21)$$

여기서,

$$\Delta w_{ji}(t+1) = \eta(x_{pi}\delta_{pj}) + \alpha \Delta w_{ji}(t) \quad (22)$$

$$\Delta w_{kj}(t+1) = \eta(o_{pj}\delta_{pk}) + \alpha \Delta w_{kj}(t) \quad (23)$$

여기서, t는 결합강도의 수정횟수, η 는 학습율, α 는 관성항계수.

여기서

$$\delta_{pk} = (t_{pk} - o_{pk})o_{pk}(1 - o_{pk}) \quad (24)$$

바이어스 θ_j , θ_k 의 수정

$$\textcircled{2} \Delta \theta_j(t+1) = \delta_{pj} + \alpha \times \Delta \theta_j(t) \quad (25)$$

$$\Delta \theta_k(t+1) = \delta_{pk} + \alpha \times \Delta \theta_k(t) \quad (26)$$

(4) 그룹핑 : 학습후 신경망 이론을 이용해서 전역조사지점을 할당하는 경우, 출력치가 최대로 되는 출력층의 유니트에 대응되는 그룹으로 입력벡터를 할당시킴.

Step 5)

- (1) $p < N$ (샘플수)이면 $p=p+1$ 로 해서 Step 2로 재이동.
- (2) $p=N$ (샘플수)이면 Step 6으로 이동.

Step 6)

- (1) $t \geq T$ 이면 학습을 종료.
- (2) $t < T$ 이면 $t=t+1$ 로 해서 Step 2로 재이동.

3. 그룹할당의 결과

판별분석과 신경망에 의한 그룹할당의 결과를 비교해보았다. 아울러 그 결과에 의한 AADT의 추정결과도 살펴보았다. 다음 <표 1>에 적응률이 나와있는데, 이는 각 조사지점이 군집분석시의 소속그룹 정보와 판별에 의해 할당된 그룹정보를 모두 갖고있기에 평가가 가능하다.

<표 1> 그룹할당의 적응률과 AADT 추정의 RMSE, U-test

그룹	지점 수	신경망 이론 이용			판별분석 이용		
		RMSE	U-Test	적응률 (%)	RMSE	U-Test	적응률 (%)
전체	171	2,046	0.0521	92.4	3,218	0.0814	32.2
그룹1	32	1,981	0.0666	93.8	3,217	0.1049	15.6
그룹2	77	2,106	0.0438	97.4	3,568	0.0740	26
그룹3	1	2,061	0.1351	0	603	0.0413	100
그룹4	13	2,222	0.0724	92.3	3,919	0.1221	46.2
그룹5	26	2,394	0.0590	92.3	2,925	0.0725	42.3
그룹6	14	994	0.1157	92.9	1,224	0.1397	28.6
그룹7	3	1,467	0.1032	0	1,281	0.0904	100
그룹8	2	529	0.1778	50	488	0.1615	100
그룹9	2	1,439	0.2482	100	1,439	0.2482	100
그룹10	1	2,520	0.1012	100	2,520	0.1012	100

<표 1>에서 대표적으로 그룹1의 결과를 살펴보면, 신경망의 경우는 적응률 92%로 상당히 높은 값이 나타났지만 판별분석의 경우는 32%로 신경망의 결과와 상당한 차이가 났다.

이 결과를 바탕으로 월변동요인과 요일변동요인을 이용한 이전 연구와 비교해 보면, 신경망 적용의 경우는 그룹할당의 적응률이 더 높아진 것으로 나타났지만, 판별분석의 경우에는 오히려 더 나빠진 것으로 나타났다.

하지만, AADT 추정의 RMSE, U-test 값들은 이전의 $\frac{1}{3}$ 수준으로 낮아져서 월변동요인만을 이용한 군집분석이 이전 방법에 비해 더 나은 결과를 갖는다고 볼 수 있었다.

이 결과는 AADT추정에 사용되는 조정계수들이 더 세분화되어 적용되었기 때문이라고 할 수 있다. 그 이유는 이전 연구에서 월그룹 3개, 요일그룹 4개로 나눠 그것만으로 모든 지점에 적용한 것에 비해, 본 연구에서는 월그룹 10개에 따른 그룹별 월별 요일그룹을 산정해, 결과적으로 계수들이 세분화되었다.

V. 결론

먼저, 그룹의 수를 변화시켜가며 각 그룹의 시간변동요인들(전체, 주중, 토요일, 일요일, 주중-토요일, 주중-일요일)의 값을 살펴보고 그 차이가 가장 뚜렷

한 변동요인으로 주중-일요일의 시간변동 요인을 선정하였다. 그런 후 월변동요인만을 사용하여 상시조사지점을 clustering하였다. 그룹간의 시간변동요인의 차이를 가장 크게 하는 것을 원칙으로 군집분석을 한 결과 10개의 그룹으로 묶을 수 있었다.

선정된 주중-일요일의 시간변동요인을 사용하여 판별분석과 신경망을 통한 그룹할당을 했다. 신경망의 적중률이 판별분석의 경우보다 훨씬 높았고, RMSE, U-test 결과도 더 좋았다.

결과를 전체적으로 살펴보면, 월변동요인만을 사용하여 군집분석한 후, 각 그룹에서 월별로 요일변동요인을 구해 적용한 AADT 추정의 결과가 이전 연구인 월변동과 요일변동을 이용한 AADT 추정의 결과와 비교하여 훨씬 실측 AADT와의 오차가 적었다. 그리고 그룹할당의 변수를 주중-일요일의 시간변동요인으로 달리하였을 때, 신경망의 경우 그룹할당의 적중률이 더 높아지는 것을 볼 수 있었다.

그러한 점에서 본 연구의 방법인 월변동요인을 이용한 clustering과 주중-일요일의 시간변동요인을 이용한 그룹할당은 AADT추정의 정확성에 있어 다수의 성과를 지닌다고 판단된다. 앞으로 기존 AADT 추정식에 다른 요인을 결합하거나 그룹할당에 쓰이는 좀

더 유용한 변수를 찾아내는 연구가 계속 되어야 할 것이다.

참고문헌

1. 김대수(1999), "신경망의 이론과 응용", 하이테크 정보세계.
2. 김정련(2000), "SAS라는 이름의 통계상자", 데이터플러스.
3. 이승재·백남철·권희정·최대순·도명식(2001), "불규칙변동 분해 시계열분석 기법을 사용한 AADT 추정", 대한교통학회지, 제19권 제6호, 대한교통학회, pp.65~73.
4. 한국건설기술연구원(2000), "연평균 일 교통량 추정을 위한 도로교통량 조사지점의 그룹핑 연구".
5. FHWA(1985, 1992, 1995, 2000), "Traffic Monitoring Guide".
6. Satish C. Sharma Pawan Lingras Fei Xu Guo X Liv(1999), "Neural networks as alternative to traditional factor approach of annual average daily traffic estimation from traffic counts", TRR1660, pp.24~31.

✉ 주 작성자 : 권희정

✉ 논문투고일 : 2001. 11. 29

논문심사일 : 2002. 3. 5 (1차)

2002. 8. 20 (2차)

2002. 10. 9 (3차)

2002. 10. 14 (4차)

2002. 11. 19 (5차)

심서관정일 : 2002. 11. 19

✉ 반론접수기간 : 2003. 4. 30