

철도수요의 시계열 분해 방법에 대한 연구

A Study on the Seasonal Decomposition of the Railway Passenger Demand

오석문*
Oh, Seog-Moon

김동희**
Kim, Dong-Hee

ABSTRACT

This paper introduces how to adopt the X-12-ARIMA to decompose the railway passenger demand of the Korea National Railroad. Especially, selecting on proper filters is focused. The trend filter is identical to the low pass filter in the signal processing field, and so the seasonal filter is to band pass filter too. Some considerations, selecting a filter, are provided from the view-point of the spectrum analysis.

The technique introduced in this paper will be adopted to the project that is to develop the forecasting system of Korea railway passenger demand which is a part of the high speed rail information system.

1. 서 론

한국철도에서 여객수요의 특징중 가장 주목할만한 사항은 명백한 요일변동이다. 철도수요의 요일변동은 각 요일별 수요량이 일정한 패턴으로 발생함을 의미한다. 철도수요의 시계열이 요일별로 일정한 패턴을 보이면, 각 요일별 지수(DOWI, Day-Of-Week Index)를 구할 수 있다. 요일지수는 시계열 조정(seasonal adjust)방법을 이용하여 구할 수 있다. 시계열 조정은 관측된 시계열(Y_t)을 추세-순환 성분(TC, trend-cycle component), 계절 성분(S_t , seasonal component) 및 불규칙 성분(I_t , irregular component)으로 나누어서 생각한다는 개념에서 시계열 분해(seasonal decomposition)라고도 한다.

미국 센서스국(CENSUS BUREAU)에서는 이러한 시계열 조정을 위해 광범위하게 적용되는 방법을 1950년대에 개발하였으며, 모형이 개량될 때마다 'X'로 시작하는 일련번호로 관리하였다. 초기에 개발된 모형을 고전적 분해(classical decomposition) 방법 또는 'CENSUS- I Method'라고 하였다. 이러한 X 계열의 모형은 처음 개발 이후 급속하게 발전하여 전세계적으로 후속모형들이 계속 발표되었으며, 그 중에서도 1960년대 미국 센서스국에서 개발한 X-11이 가장 눈부신 발전이었다. X-11에서 사용하는 핵심적인 필터(basic filter)들은 현재 주요 통계 패키지(SAS, AREMOS, STATISTICA 등)에 포함되어 있다. 1980년에 캐나다 통계청에서는 X-11-ARIMA를 발표하였다. 이후 1988년에는 X-11-ARIMA/88가 발표되었는데, 이것은 X-11-ARIMA가 88년도에 크게 개선되었다는 의미이다. 현재에는 X-12-ARIMA까지 개발된 상태인데, 이것은 미국 센서스국에서 1990년대 후반 개발한 것이다¹⁾²⁾. X-11/X-12 모형은 OECD, IMF 등에서 적용하고 있다고 보고되고 있으며³⁾, Lee, Geung-Hee(2001)⁴⁾는 한국은행에서 BOK(Bank Of Korea)-X-12-ARIMA를 개발하여 활용하는 사례를 발표하였다. X-11/X-12에 상응하는 프로그램으로 스페인 은행의 TRAMO/SEATS과 독일의 BV4도 같은 맥락의 프로그램이다.

그러나 이러한 시계열 분해 모형의 핵심 필터들은 모두 12개월 주기(monthly) 또 4분기 주기(quarterly)의 관측값을 처리하도록 설계되어있다(예 X-11에서 2×12 MA, Spencer's 15 Weighted MA 등). 이것은 이러한 모형을 개발한 주체들이 모두 경제지표나 고용지표에 대한 전환점(turning point)을 정확하게 알고 싶어하는 목적이 컸기 때문이다.

본 논문에서는 X-12-ARIMA의 절차에 따라 철도수요의 시계열을 분해하고, 분해된 계절 성분으로부터 요일지수를 결정하는 방법을 소개한다. X-12-ARIMA에 의한 시계열 조정은 12개월을 주기로 한 방법이다. 12개월을 1주기로 하는 경우와 7일을 1주기로 하는 방법에서는 조정과정 중에 적용하는 이동평균(Moving Average) 방법이 상이하다. 특히 X-12-ARIMA에서는 12개월을 주기로 하는 헨더슨의 가중이동평균(Henderson's Weighted Moving Average)을 적용하고 있다. 본 논문에서는 이러한 주기성(periodicity)의 변

* 한국철도기술연구원 주임연구원, 정회원

** 한국철도기술연구원 선임연구원, 정회원

화에 따른 추세 필터(trend filter)의 선택방법에 대해 설명한다. 마지막으로 한국철도의 여객수요(예, 서울~부산, 1998.2.1~1998.9.30) 시계열 자료를 X-12-ARIMA의 절차에 따라 분해하고, 요일지수를 계산한 예를 제시한다. 시계열 분석의 업무에 X-12-ARIMA를 이용하면 시계열을 추세-순환, 계절, 불규칙 성분으로 정량화할 수 있고, 계절 성분에서 요일지수를 정량화하여 수송계획에 대한 기초 지식이 될 수 있다.

2. 본 문

2.1 X-12-ARIMA의 절차에 따른 철도수요의 시계열 조정

그림 1은 본 논문에서 샘플 자료로 사용하는 시계열 관측값(Y_t)과 이에 대한 스펙트럼 분석 결과(Y_f)를 나타내고 있다. 그림으로부터 철도수요가 추세 성분, 계절 성분 및 불규칙 성분으로 구성되어 있음을 직관적으로 알 수 있다.

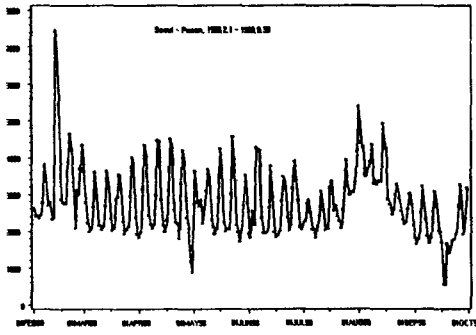


그림 1-A 시계열 관측값(Y_t)

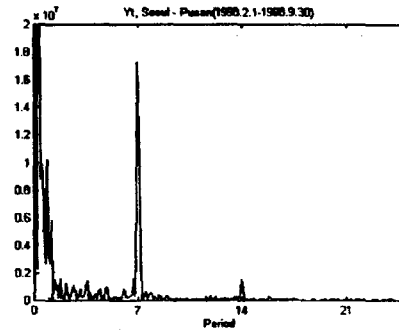


그림 1-B 관측값 스펙트럼(Y_f)

2월중의 피크는 설날 연휴에 대한 대수송의 결과이며, 이러한 성분에 대한 처리는 일별(daily) 자료를 다루는 연구자들에게는 별도의 연구테마가 된다. 특히 우리나라와 같이 음력의 절기에 의한 공휴일의 위치가 양력에 대해 매년 변동함으로 인해 발생하는 불규칙 성분 때문에 일반적으로 적용할 수 있는 일별 모형을 개발하는 것은 매우 어려운 환경이다. 따라서 일별 자료의 특성 때문에 나타나는 불규칙성 및 극단값(outlier)의 문제는 본 논문에서는 다루지 않는다.

8, 9월 부분의 관측값에는 비교적 강한 추세변동이 나타나고 있다. 그러나 무엇보다 명백한 것은 매 1주일을 주기로 반복되는 패턴으로 관측값의 스펙트럼 결과를 보면 알 수 있다. 이 절에서는 철도수요 시계열의 분해절차를 X-12-ARIMA의 전체적인 절차에 따라 설명하고, 주기성의 차이에 의해 발생하는 문제에 대해서는 주어진 시계열에 적합한 필터를 선택하는 방법을 제시하면서 전체적인 계열의 조정과정을 보이고자 한다.

시계열 자료를 분해할 때, 시계열과 각 성분은 다음과 같은 관계를 가질 수 있다.

$$Y_t = TC_t \times S_t \times I_t \quad (1-1)$$

$$Y_t = TC_t + S_t + I_t \quad (1-2)$$

곱셈모형(1-1)은 시간의 변화에 따라 계절성분의 진폭이 크게 변화할 경우에 적합하고, 그렇지 않을 경우에는 덧셈모형(1-2)이 적합하다고 알려져 있다.

시계열 자료가 준비되면 먼저 그 계열의 초기 추세-순환계열 성분(TC_t^i , initial trend-cycle)을 구해야 한다. 초기 추세-순환 필터로 X-11/X-12에서는 2×12 MA 및 2×4 MA가 사용되고 있다. 그러나 이러한 필터들은 12개월 또는 4분기를 주기로 하는 자료를 대상으로 하기 때문에 철도수요와 같이 7일 주기의 자료에 대해 적용할 경우에는 원자료를 지나치게 또는 부족하게 완화

(over/under smoothing)하여, 다음의 요일성분 추출에 영향을 줄 우려가 있다. 주어진 자료에서는 주기가 7이고, 그 자체가 홀수이므로 7 MA를 적용하여 추세-순환 성분의 계열을 구할 수 있다. 추세-순환 성분이 구해지면 식(2-1)~(2-5)와 같이 예비적 시계열 조정을 시행할 수 있다.

$$TC_1^1 = 7 MA(Y_t) \quad (2-1)$$

$$Y_t/TC_1^1 = S_t \cdot I_t^1 \quad (2-2)$$

$$\widehat{S}_t^1 = 3 \times 3 MA(S_t \cdot I_t^1) \quad (2-3)$$

$$S_t^1 = \text{Normalizing}(\widehat{S}_t^1) \quad (2-4)$$

$$Y_t/S_t^1 = TC_1^1 \cdot I_t = A_t^1 \quad (2-5)$$

식(2-1)과 (3-1)의 차이에 대해서는 필터선정에 관한 문제로 '2.2'에서 자세히 설명하기로 한다. (2-2)의 결과에서는 요일(계절) 성분과 불규칙 성분이 공존하고 있다. 식(2-2)의 불규칙 성분은 (2-3)과 같이 기대값(expectation)을 취하면 제거된다. 결과적으로 (2-4)와 같이 예비적인 요일성분을 구할 수 있다. 예비적인 시계열 조정의 결과는 다시 중간단계의 시계열 조정과정을 거친다.

$$TC_2^2 = \text{Henderson's MA}(A_t^1) \quad (3-1)$$

$$Y_t/TC_2^2 = S_t \cdot I_t^2 \quad (3-2)$$

$$\widehat{S}_t^2 = 3 \times 5 MA(S_t \cdot I_t^2) \quad (3-3)$$

$$S_t^2 = \text{Normalizing}(\widehat{S}_t^2) \quad (3-4)$$

$$Y_t/S_t^2 = TC_2^2 \cdot I_t = A_t^2 \quad (3-5)$$

$$A_t^2/TC_2^2 = I_t^2 \quad (3-6)$$

예비적으로 조정된 시계열(A_t^1)에 Henderson의 가중 이동평균을 취하면 개정된 추세를 구할 수 있다. X-12-ARIMA의 절차에서는 이 단계에서 23-term Henderson MA가 적합하다고 보고되었다⁵⁾. (3-2), (3-3) 및 (3-4)의 절차는 (2-1), (2-2) 및 (2-3)의 절차와 같은 방법으로 구하고, 그 결과에 따라 중간단계의 조정 및 불규칙 성분을 (3-5)와 (3-6)과 같이 구할 수 있다. 중간단계의 시계열 조정과정은 식(4-1)과 (4-2)와 같이 최종 조정된다.

$$TC_3^3 = \text{Henderson's MA}(A_t^2) \quad (4-1)$$

$$A_t^2/TC_3^3 = I_t^3 \quad (4-2)$$

2.2 주기 7의 시계열을 위한 추세필터 선정

주기가 7인 철도수요의 추세필터 선정절차는 매우 정교한 절차를 거쳐야 한다⁶⁾⁷⁾⁸⁾. X-12-ARIMA를 포함하여 시계열 조정의 절차에 사용되는 추세필터는 모두 MA 개념의 필터가 사용된다. 따라서 MA의 항(term) 수에 따라서 필터의 결과가 달라진다. 항 수가 지나치게 큰 경우(over smoothing) 자료의 요일성(week periodicity)이 손상을 받을 우려가 있고, 또 항 수가 너무 작은 경우(under smoothing)에는 추세성분에 포함될 수 있다. 그림 2-A는 7 MA(Y_t)의 결과이고, 그림 2-B는 23/7/3 MA(Y_t)에 대한 스펙트럼 분석의 결과이다. 그림에서 보는 바와 같이 23 MA(Y_t) 필터는 차단 주파수(Fc, cut-off frequency)가 7 주기 이하에서 형성되고 있고, 3 MA(Y_t) 필터는 7 주기 이하에서 주기적인 성분이 포함되어 있다.

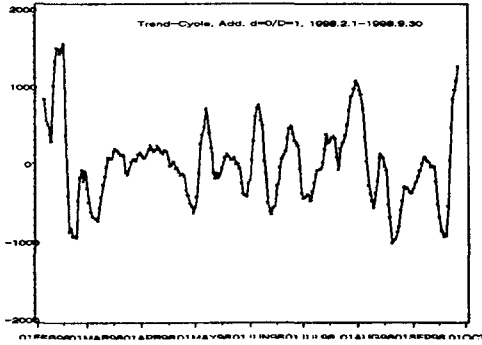


그림 2-A 예비추세 (TC_i)

그림 2-B 추세 필터(3/7/23 MA)

따라서, 시계열 조정을 위한 추세필터를 선정할 때는 반복되는 주기에서 차단 주파수가 형성되는 가장 작은 항 수의 MA 필터를 선정해야 한다. 이것은 신호처리(signal processing) 분야에서 저역통과 필터(low pass filter) 설계의 개념과 상통한다.

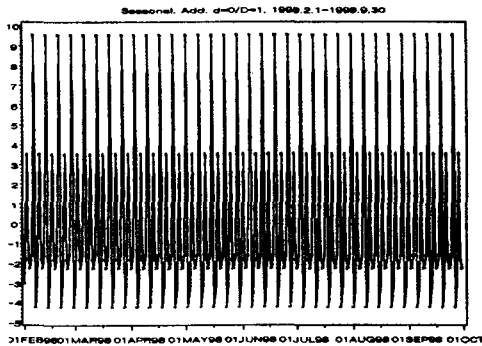


그림 3-A 요일성분(S_i)

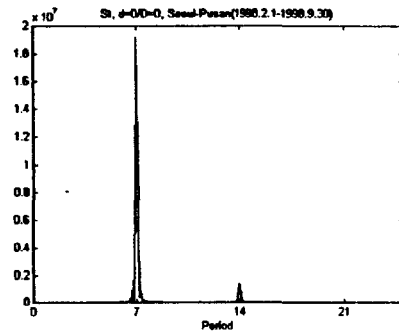


그림 3-B 요일성분(S_i)

주어진 계열이 추세 필터를 거치고 나면 다시 요일 필터를 거치게 되는데, 그림 3-A, 3-B는 요일 필터를 거쳐서 추출된 성분이다. 그림에서 보는 보와 같이 7 주기의 성분만이 추출되었으며, 이것은 신호처리 분야에서 대역통과 필터(band pass filter) 설계의 개념과 상통한다.

2.3 철도수요의 요일지수

식(2-2), (3-2)와 같이 추세 성분이 조정된 계열은 요일 성분과 불규칙 성분만 남는 추세조정 계열이 된다. 추세조정 계열은 일정한 주기로 반복되는 성분이 지배적이므로, 전체 계열은 유사한 패턴을 가진 소계열의 반복이라고 생각할 수 있다.

$$Y_i \sim M(S_i, \sigma_i^2), \quad i = 1, 2, \dots, 7 \quad (5-1)$$

$$S_i = 3 \times 5 \text{ MA}(Y) \quad (5-2)$$

여기서 i 는 i 번째 요일(i -th DOW)을 가리키는 식별자이고, Y_i 는 i 번째 요일의 추세조정 계열 값이 된다. Y_i 는 이미 추세조정을 거친 계열이므로, 자기상관계수(ACF, autocorrelation factor)가 '0'에 근접해야 하고, σ_i^2 는 유의한 수준으로 충분히 작아야 하며, 각 i 에 대해 독립적이어야 한다.

식(5-2)는 식(2-4), (3-4)와 같이 정규화 과정을 거쳐서 요일지수를 구하게 된다.

표 1 샘플 자료에 대한 요일지수

DOW	일	월	화	수	목	금	토
INDEX	0.279	-0.161	-0.282	-0.311	-0.246	0.195	0.526

요일지수는 덧셈모형인 경우 $\sum S_i = 0$ 과 곱셈모형인 경우 $\sum S_i = 1$ 의 특징을 갖는다. 표 1은 샘플 자료의 요일지수를 구한 것이다.

3. 결 론

본 논문에서는 한국철도 여객수요의 시계열을 성분별로 분해하는 시계열 분해방법을 소개하였다. 특히 경제 및 노동분야에서 널리 이용되는 X-12-ARIMA(12/4-term)를 철도수요에 적용하는 방법에 대해 설명하였다. X-12-ARIMA를 철도에 적용하기 위해서는 별도의 추세필터를 선택해야 하는데, 이에 대한 선정방법에 대해 설명하였다. 다음으로 추세조정된 계열로부터 요일지수를 구하는 방법에 대해 설명하고 예를 제시하였다.

시계열 조정 방법은 철도수요의 가장 큰 특징인 요일변동을 정량적으로 설명할 수 있는 방법으로 철도수요를 정확하게 이해하는 기초지식을 제공할 것이다. 철도수요에 대한 기초지식은 차후 수익성있는 수송계획을 작성하기 위한 기초자료가 된다.

하지만 일별 자료의 분석에서는 요일지수를 구하는 것뿐만 아니라 극단값의 문제, 음력에 의한 공휴일의 이동문제 등 매우 복잡한 문제가 포함되어 있다. 이러한 부분에 대한 해결방법으로 365일 전체에 대한 일별지수를 모두 구하는(Young J, Joo, 1996⁹⁾) 시도가 있었다. 그러나 철도에서는 OD의 수가 수천에 달하므로 이러한 방법을 적용하기에는 무리가 있을 것으로 보인다. 따라서 철도수요와 같이 수많은 계열을 한꺼번에 고려해야 하는 경우, 지배적인 패턴인 요일변동에 대해 조정 후, 일별자료의 특성에 따른 문제는 전문가가 관여하는 방법을 따라야 한다.

본 논문에서 소개한 방법은 고속철도 통합정보 시스템의 수요예측 시스템 개발을 위한 분석단계에 적용하여 철도수요의 요일별 특징 분석에 활용할 예정이다.

감사의 글

이 논문의 준비과정에서 게시판(<http://forecast.kaist.ac.kr>) 토론에 참여해 주신 전덕번 교수님이 하 대학원생들과 이메일로 도움을 주신 Bureau of the Census, Statistical Research Division의 David F. Findely께 감사의 뜻을 표합니다.

참고문헌

- 1) Time Series Staff, Statistical Research Division of U.S. Census Bureau, X-12-ARIMA Reference Manual Version, Jan 17, 2001.
- 2) David F. Findely, Brian C. Monsell, William R. Bell, Mark C. Otto and Bor-Chung Chen, "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program", Journal of Business and Econometric Statistics, vol. 16(1998), no. 2
- 3) 김연형, 시계열 분석과 예측, 자유아카데미, p.284
- 4) Lee, Gung-Hee, Hwang, Hee-Jin and Chang, Young-Jae(2001), "X-12-ARIMA Seasonal Adjustment in Korean National Income Statistics and Money and Banking Statistics", 52th ISI Seoul Forum Proceeding

- 5) Masahiro Higo and Sachiko Kuroda Nakada, "How Can We Extract a Fundamental Trend from an Economic Time Series ?", *Monetary and Economics Studies*, Dec. 1998
- 6) Alistair Gray and Peter Thomson, "Design of Moving-Average Trend Filters using Fidelity, Smoothness and Minimum Revisions Criteria"(1996), Bureau of the Census, Statistical Research Division, Statistical Research Report Series No. RR96/01.
- 7) Alistair Gray, "How the X-11 program implements a trend-seasonal decomposition"(1990), working paper.
- 8) William Bell and Brian Monsell, "X-11 Symmetric Linear Filters and their Transfer Functions"(1992), Bureau of the Census, Statistical Research Division Research Report Series No. RR-92/15
- 9) Young J, Joo and DUK BIN JUN(1996), "Forecasting A Daily Time Series With Varying Seasonalities: An Application to Daily Visitors to Farmland in Korea", *Computers ind. Engng.* vol. 30, no. 3 pp. 365 ~ 373