

A Refinement of Point Forecast Using Dependency Structure in Irregular Component of BOK-X12-ARIMA¹⁾

S. Y. Hwang²⁾ · S. K. Yang³⁾

abstract

BOK-X12-ARIMA has been developed by the Bank of Korea in order to accommodate special features such as lunar effect, labor day and election effect which are intrinsic in Korean seasonal time series. Irregular component resulting from BOK-X12-ARIMA is usually treated as white noise time series. If this shows dependency structure, it may be advisable to incorporate dependency in irregular component into prediction. This article illustrates how to refine point forecast using dependency structure in irregular component.

Keywords : BOK-X12-ARIMA, Forecast, Irregular component

1. 서 론

계절형 시계열 분석의 가장 큰 목적은 과거의 패턴이 지속적으로 유지된다는 가정 하에서 현재까지 수집된 자료들을 분석하여 미래에 대한 예측을 하는 것이다. 미래의 계획을 합리적으로 설계하고 이에 따른 리스크를 축소하기 위해서는 정확한 예측이 필요하다. 즉, 정확한 예측이 이루어져야 현재 합리적인 의사결정이 가능해지는 것이다.

기존의 계절형 시계열의 분석 및 예측 방법으로 지수평활법, Winters방법 등 고전적인 방법과 확률적 시계열 모형인 Seasonal-ARIMA모형 등이 이용되어 왔다. 또 관측된 시계열을 변동요인별로 즉, 추세요인(Tt), 순환요인(Ct), 계절요인(St) 등으로 분해하여 예측하는 방법(분해법)도 있다. 분해법 중 미 상무부가 개발한 X12-ARIMA법

1) This work was supported by a grant(2005) from Sookmyung Women's Univ.

2) First Author : Professor, Department of Statistics, Sookmyung Women's Univ., Seoul, 140-742, Korea.
E-mail : shwang@sookmyung.ac.kr

3) Graduate student, Department of Statistics, Sookmyung Women's Univ., Seoul, Korea.

을 음력을 바탕으로 한 설, 추석 및 공휴일, 불규칙적으로 시행된 선거 및 각종 정책 등에 영향을 받는 우리나라의 특성에 맞게 조정한 한국은행이 개발한 BOK-X12-ARIMA 방법이 우수한 것으로 알려져 있다(cf. 이궁희(1997,1998,1999,2003)).

각 분석방법을 통해 얻어진 모형의 잔차는 기본적으로 백색잡음을 가정하고 있는데 만일 잔차(불규칙 요인)에서 종속구조가 통계적으로 의미 있게 나타나는 경우에는 이를 모형화 시켜 보정(adjustment) 단계를 걸쳐 점 예측치의 예측력을 향상시킬 필요를 느끼게 된다. 최근에 Hwang and Yang(2005)는 다양한 계절 시계열 분석 기법들의 불규칙요인을 변동성 모형으로 적합 시켰으며 이를 통해 얻어진 조건부 분산(conditinal variance)의 변화에 따라 시계열 예측구간을 융통성 있게 제시한 바 있다. 본 논문에서는 BOK-X12-ARIMA방법의 불규칙 요인을 이용하여 점 예측치 개선방안을 제시하고 이를 국내 시계열 자료에 적용시켜 보고자 한다.

2. 한국형 BOK-X12-ARIMA

월별 또는 분기별 경제시계열은 계절적으로 1년을 주기로 반복해서 변동하는 성분을 포함하는 경우가 많은데, 이러한 변동을 계절성 또는 계절변동성분이라 한다. 계절성이 발생하는 요인으로서는 먼저 기후 등 자연조건의 변화를 들 수 있으며 명절, 축제일 등 사회적 관습과 경영환경에 관한 제도 등도 이에 영향을 준다. 월별 또는 분기별 시계열은 일반적으로 추세변동(T_t), 순환변동(C_t), 계절변동(S_t), 불규칙변동(I_t)으로 구성되어 있다고 가정한다. 우리나라 경제시계열을 분석하기 위해서는 좀더 세분하여 계절변동(S_t), 명절성분(H_t), 요일구성성분(D_t), 추세변동성분(T_t), 순환변동성분(C_t) 및 불규칙변동성분(I_t)으로 구성되어 있는 것으로 가정하기도 한다. 여기서 명절변동은 음력에 바탕으로 한 설과 추석에 따른 변동이며, 요일구성변동은 월 또는 분기별 요일구성에 따른 변동을 의미한다. 계절조정이란 1년 주기의 계절변동성분은 물론 명절, 요일 등 규칙적 변동을 통계적으로 추출하여 원계열로부터 제거하는 절차를 의미한다. 좀더 자세한 내용은 김수아(1998)와 이궁희(1998)를 참고하기 바란다.

계절조정 및 계절조정을 위한 통계소프트웨어에 대한 연구는 세계 각국에서 활발하게 이루어지고 있으며 X11법, X11-ARIMA법, X11-ARIMA/88, X12-ARIMA법 베타 버전에 이어 X12-ARIMA법 버전 0.1이 발표되었다(Findley et al(1998)). 미 상무성에서 1996년 초 그 동안의 계절조정에 대한 연구 성과를 바탕으로 이동평균법에 모형접근법을 가미한 새로운 계절조정방법 X12-ARIMA방법을 제안하였다. X12-ARIMA 방법은 회귀분석 기법을 이용하여 원통계를 사전조정된 후 이동평균방법을 적용하는 방법인데 크게 사전조정부분, 이동평균부분, 사후진단부분의 3개 부분으로 구성되어 있다.

먼저, 사전조정부분에서는 RegARIMA(Regression과 ARIMA의 합성어로 더미변수(dummy variable)에 의한 회귀모형과 ARIMA 모형의 결합을 의미함)를 이용하여 구조변화, 이상치 및 명절변동요인 등을 추정하여 원계열을 사전조정하고 이를 이용하

여 예측치를 ARIMA로 연장, 사전조정원계열을 작성한다. 둘째, 이동평균부분에서는 앞서 사전조정원계열에 대하여 X11법을 이용하여 시계열을 분해하여 계절조정계열을 생성한다. 마지막으로 사후진단부분에서는 계절조정의 적절성 평가를 위하여 스펙트럼 및 Sliding span 분석 등을 실시하여 불규칙변동에 계절변동이 남아있는지, 구간을 달리해도 계절변동조정이 안정적인지 진단한다. 검정결과 적절하지 않은 것으로 평가 되면 다시 조정하여 앞의 3단계를 반복한다. 보다 자세한 내용은 Findly et al.(1998), Bureau of the Census (1998)등을 참조하기 바란다.

한국형 BOK-X12-ARIMA :

X12-ARIMA는 지금까지의 계절조정연구 및 각국 통계기관의 요구를 반영하여 작성되었다. 그러나 우리나라의 경제시계열은 음력을 바탕으로 한 설, 추석 및 공휴일, 불규칙적으로 시행된 선거 및 각종 정책 등에 영향을 받고 있기 때문에 X12-ARIMA를 우리 시계열의 특성이 반영되는 조정할 필요가 있다. 한국형 BOK-X12-ARIMA법은 우리 통계 특성이 제대로 반영되도록 다음과 같은 조정과정을 반영하였다(이궁희(1997,1998)).

1. RegARIMA모형을 이용한 (태음력)명절효과의 조정
2. RegARIMA모형을 이용한 영업일수 및 요일변동의 조정
3. 특이항 및 구조변화의 조정
4. 캐나다 통계청의 5개 예측모형이 우리사정에 맞지 않는 점을 고려하여 표준모형을 28개로 확장. 예측모형에서 이용되는 5개의 표준모형은 캐나다 자료를 바탕으로 작성되었기 때문에 변동성이 높은 우리 시계열에 대해서 이를 그대로 적용하는 데에는 한계가 있으며 따라서 우리 경제시계열에 적절한 표준모형의 선정이 필요하며 박유성과 최윤희(1997)는 예측력(MAPE)를 기준으로 국내 102개 통계청 생산관련 시계열에 대해 $(0 \ 1 \ 1)(0 \ 1 \ 1)_4$, $\log(1 \ 1 \ 0)(0 \ 1 \ 1)_4$, $\log(0 \ 1 \ 1)(0 \ 1 \ 1)_4$, $\log(0 \ 1 \ 1)(2 \ 1 \ 0)_4$, $\log(0 \ 1 \ 2)(0 \ 1 \ 1)_4$, $\log(1 \ 1 \ 0)(0 \ 1 \ 2)_4$ 의 6개 표준모형을 제시한 바 있다.

3. 불규칙요인의 종속구조를 활용한 예측치의 보정

불규칙요인이 백색잡음이 아니라는 확신이 있다면 불규칙요인의 종속구조로부터 의미 있는 부분을 추출해 내고 이를 이용하여 계절시계열 예측에 대한 점 예측치(point forecast)와 95%-예측구간(forecast interval)의 개선방안을 고려할 수 있다. 여기서 예측으로서 1시차 후 예측을 고려하도록 한다. 관측 시계열 $\{y_t\}$ 의 조건부 기대값을 $\{M_t\}$ 라 할 때 y_t 의 최소평균제곱오차(MMSE) 기준에서의 최적(optimal) 1시차 후 예측치는 M_t 이다(Chan(2002)). 단기예측에 따른 예측오차를 ϵ_t 로 표시할 때 다음 식이 성립한다.

$$y_t = M_t + \epsilon_t$$

여기서 예측오차 ϵ_t 는 분해시계열에서의 불규칙 요인인 I_t 와 유사하며 M_t 는 분해시계열의 주요 신호(signal)인 추세, 순환, 계절요인을 나타낸다고 할 수 있다.

3.1 점 예측치의 보정

시점 n 까지의 자료 y_1, y_2, \dots, y_n 에 근거한 1시차 후 예측치를 $y_n(1)$, BOK-X12-ARIMA 분석기법으로부터 얻은 점 예측치를 $M_n(1)$ 이라 할 때 예측오차 ϵ_t 이 백색잡음이라면 $y_n(1) = M_n(1)$ 이 성립하지만 예측오차 ϵ_t 이 백색잡음이 아니라는 확신이 있다면 예측오차 ϵ_t 의 실현값인 잔차를 이용하여 보정해 줄 필요가 있게 된다. 보정항을 ϵ_t^* 라 하면, ϵ_t^* 은 적절한 시계열 모형으로부터 계산된다. 따라서 1시차 후 예측치로서 다음과 같은 식을 제안한다.

$$y_n(1) = M_n(1) + \epsilon_{n+1}^* \quad (1)$$

3.2 구간예측치의 보정

예측치 y_{n+1} 에 대한 신뢰수준 95%의 1-시차 후 예측구간은 다음과 같다.

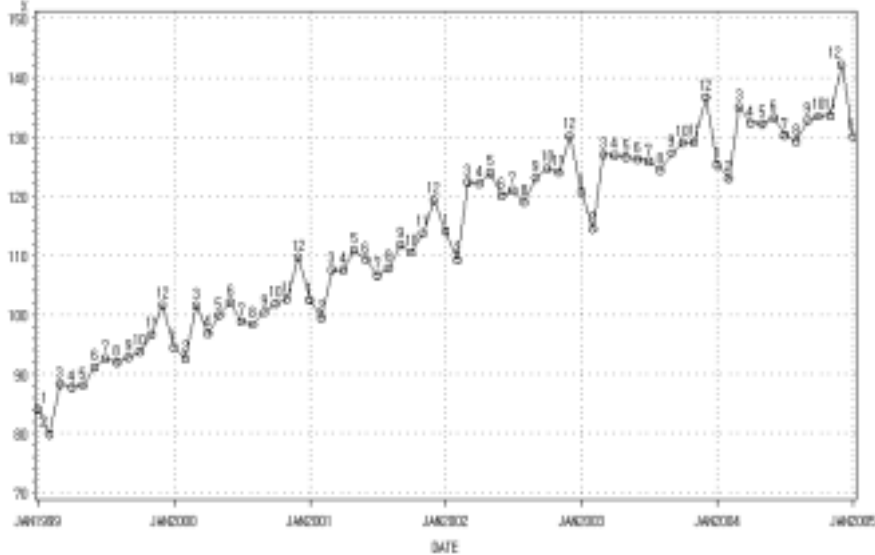
$$y_n(1) \pm 2\sqrt{h_n}, \quad h_t = \text{Var}(\epsilon_t | \Psi_{t-1})$$

여기서 $y_n(1)$ 는 식(1)을 통해 보정된 점 예측치이다. 변동성 모형(예를 들어 ARCH 모형)을 고려하지 않은 기존 방법으로부터의 95% 예측구간의 너비(W)는 $W = 4\sqrt{h_t}$ 이다. 변동성 모형을 고려하지 않은 경우에는 예측구간의 너비 W 는 시점에 관계없이 일정(time homogeneous)하다. 하지만 변동성이 시점에 따라 변화하는 자료의 경우에는 변동성 모형으로부터 잔차(불규칙요인)의 조건부 분산식인 h_t 를 유도하여 h_t 의 변화가 심한 경우에는 h_t 의 크기에 따라 예측구간의 너비를 보정해 줄 필요가 있게 된다. 자세한 내용은 Hwang and Yang(2005)를 참고하기 바란다.

4. 사례분석

본 절에서는 1999년 1월부터 2004년 12월까지의 우리나라 (월별)서비스업활동지수 (통계청-도소매업, 서비스업-서비스업활동지수(서비스업동태조사))를 이용하여 앞장에서 언급된 방법론을 예시하도록 한다. 다양한 계절시계열 분석법들과 비교하기 위해 1999년 1월부터 2004년 12월까지의 자료는 예측을 위한 분석에 사용하고 2005년 1월의 자료는 표본외 자료로 두어 각 방법들에 의해 얻어지는 예측값과 실제의 자료를 가지고 예측력을 판단하고자 한다. 지수평활법, 승법윈터스방법, 가법윈터스방법은 SAS/ETS/Proc FORECAST를 이용하였고, Seasonal ARIMA는 Proc ARIMA를 그리고 계절조정 프로그램은 BOK-X12-ARIMA를 사용하였다. 또 각 분석방법의 MSE, MAPE를 구하여 적합도를 비교하였다.

자료의 분석에 들어가기 전에 앞서 시계열자료의 특성을 알아보기 위해 시도표를 그려보면 <그림 1>과 같다.



<그림 1> 서비스업활동지수 시도표

서비스업활동지수의 seasonal ARIMA 모형은 $(4, 1, 1) \times (0, 1, 0)_{12}$ 로서 최대우도법(ML) 추정 결과는 다음과 같다.

$$w_t = (1 - B)(1 - B^{12})y_t,$$

$$(1 + 0.885B + 0.378B^2 - 0.358B^4 + 0.283B^{10})w_t = (1 + 0.553B)\epsilon_t$$

BOK-X12-ARIMA를 이용하여 계절조정을 한 결과는 다음과 같다.

<표 1> BOK-X12-ARIMA분석 결과

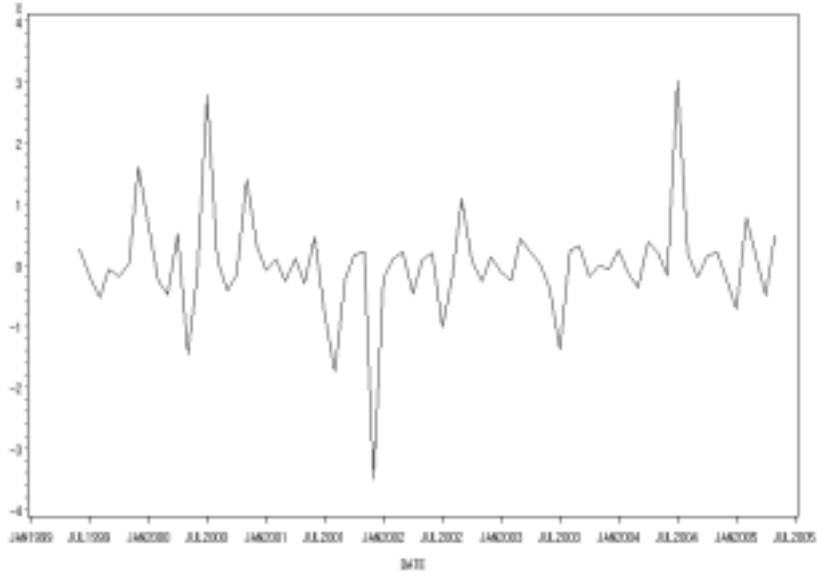
계절성 존재 여부			명절 변동				공휴일 (WD)	요일 구성	ARIMA 모형	안전성 분석	
안정적 계절성	이동 계절성	안정적 계절성 여부	실		추석					Q 통계량	Silding span분석
			전	후	전	후					
○	○	○	1.33	-0.61	0.00	-1.69	-2.92	0.00	(011)(011)	0.33	안정적

다양한 분석 방법의 MSE와 MAPE는 다음과 같다.

<표 2> 서비스업활동지수 각 분석방법의 MSE, MAPE

	지수평활법	승법윈터스 방법	가법윈터스 방법	계절형 ARIMA	BOK-X12-ARIMA
MAPE	2.7240	2.2558	2.1108	1.2020	0.4237
MSE	16.4555	11.1205	9.8438	3.3096	0.8346

<표 2>에서 보면 BOK-X12-ARIMA의 적합력이 가장 우수하며, 그 다음으로 seasonal ARIMA가 우수함을 알 수 있다. 적합력이 가장 우수한 BOK-X12-ARIMA 방법에 대하여 예측치 보정에 대하여 살펴보도록 하자. <그림 2>는 서비스업활동지수를 BOK-X12-ARIMA로 적합 시킨 결과 얻은 불규칙요인 그래프이다.



<그림 2> (서비스업활동지수) BOK-X12-ARIMA방법의 잔차 플롯

이를 살펴보면 변동성이 약간 존재하는 것으로 보이지만 정확한 검정을 위해 SAS/ETS/Proc autoreg의 ARCH test를 수행해본 결과 조건부 이분산은 나타나지 않았다. backstep 옵션을 수행해 본 결과 시차 13, 31이 선택되었으며 불규칙요인(ν_t)을 AR(13, 31)모형을 적합 시킨 결과 다음 식을 얻었다.

$$\nu_t = -0.2212\nu_{t-13} - 0.5935\nu_{t-31} + \varepsilon_t$$

불규칙요인을 AR 모형으로 적합 시킨 결과를 이용하여 2005년 1월 서비스업활동지수의 ε_{n+1}^* 을 구한 후 식(1)을 이용하여 보정된 점 예측치를 구해 보았다. 그 결과는 다음의 <표 3>과 같으며 BOK-X12-ARIMA 법이 조정 전 130.147에서 129.909로 실제 값(130)에 가장 가깝게 개선된 것을 확인할 수 있다.

<표 3> 보정된 점 예측치(실제 2005년 1월의 서비스업활동지수 = 130)

	지수평활법	승법윈터스 방법	가법윈터스 방법	계절ARIMA 방법	BOK-X12-ARIMA
보정 전	136.686	130.929	131.779	131.150	130.147
보정 후	141.135	130.649	132.259	131.150	129.909

참고문헌

1. Bureau of the Census (1998). *X-12 ARIMA Reference Manual*. Final Version 0.1.
2. Chan, N. H.(2002). *Time Series Applications to Finance*. Wiley.
3. Findly, D. F., Brian C. Monsell, William R. Bell, Mark C. Otto and B.C. Chen (1998). New capabilities and methods of the X-12 ARIMA seasonal adjustment program, *Journal of Business and Economic Statistics*, 16, No. 2.
4. Lee, G.H. (2003), Korean traditional holiday adjustment with RegARIMA, ISI Proceedings.
5. 김수아 (1998). 국내 계절시계열 자료 분석 : X-12-ARIMA를 중심으로. 숙명여자대학교 통계학과 석사학위논문.
6. 박유성 · 최연희 (1997). 한국형 X-11 ARIMA 프로시저에 관한 연구. 춘계 학술발표회 논문집, 한국통계학회.
7. 이금희 (1997). X-12 ARIMA방법에 의한 국민소득 및 통화통계에 대한 계절조정 시산결과.
8. 이금희 (1998). 한국경제시계열의 계절조정 방법 : X-12 ARIMA 법을 중심으로. 경제분석, 4권 1호, 한국은행 금융경제연구소.
9. 이금희(1999). 한국형 계절변동조정 프로그램 개편 : BOK-X-12-ARIMA.
10. 황선영 · 양수경(2005). 월별 시계열 자료 분석기법들의 비교분석 및 변동성 모형을 이용한 단기예측력 개선 방안. *Quarterly National Accounts*, 23권 4호, 33-68, 한국은행 경제통계국.

[2005년 12월 접수, 2006년 2월 채택]