信頼性應用硏究
제3권, 제1호, pp. 13-40, 2003

전자제품 판매매출액 시계열의 계절 조정과 수요예측에 관한 연구

서 명 율, 이 종 태

동국대 산업시스템공학부

A Study on the Seasonal Adjustment of Time Series and Demand Forecasting for Electronic Product Sales

Myeong Yul Seo, Jong Tae Rhee

Department of Industrial & System Engineering Dongguk University

Abstract

The seasonal adjustment is an essential process in analyzing the time series of economy and business. One of the powerful adjustment methods is X11-ARIMA Model which is popularly used in Korea. This method was delivered from Canada. However, this model has been developed to be appropriate for Canadian and American environment. Therefore, we need to review whether the X11-ARIMA Model could be used properly in Korea. In this study, we have applied the method to the annual sales of refrigerator sales in A electronic company. We appreciated the adjustment by result

analyzing the time series components such as seasonal component, trend-cycle component, and irregular component, with the proposed method. Additionally, in order to improve the result of seasonal adjusted time series, we suggest the demand forecasting method base on autocorrelation and seasonality with the X11-ARIMA PROC.
Keyword : Seasonal Adjustment Time Series, X11-ARIMA, Demand Forecasting.

1. 서론

초기단계의 시계열분석은 외삽법이나 평활법 등으로 순환변동 또는 계절변동 등의 내용을 파악한 다음 미래예측을 수행하는 단계에 머물렀으나 이제는 Box-Jenkins에 의해 제시된 ARIMA 모형을 이용하여 불규칙 변동까지도 고려하는 보다 정밀한 방법론이 일반적으로 사용되고 있다. 경영 및 공학에서 나타나는 시계열 자료는 대부분 계절적요소, 추세적요소 그리고 랜덤요소로 구성된다. 이러한 세가지 요소를 포괄하는 일반적인 형태로 SARMA(Seasonal Autoregressive Moving Average) 모형[15]이 있으나 분석과정에 전문가적인 판단이 요구되기 때문에 실용화되지 못하고 있다. 계절적 요소를 제외하고 나머지 요소를 고려하는 경우ARMA(Autoregressive Moving Average)모형이 된다. ARMA 모형을 전문지식이 없이 분석할 수 있게 하는 기법으로, AIC(Akaike's Information Criterion)을 이용한 방법[14], DDS(Data Dependent System[24], Pade Table을 이용한 방법, VSACF(Vector Sample Autocorrelation Function)을 이용한 방법[11]등이 개발되었다. 이러한 기법들중에서 VSACF 방법이 통계적 특성 및 계산에 있어서 효과적이지만, 계절적 요소가 존재하는 시계열에 적응하기 위해서는 시계열로부터 계절적 요소를 분리하는(Deseasonlize)과정이 선행되어야 한다[11]. 계절적 요소가 존재하는 시계열의 분석방법으로는 평활 모형을 사용하는 Winter's 방법[26]과 분활기법을 사용하는 CensusⅡ[22]를 들 수 있다. 그러나 Winter's 방법은 MSE를 최소화하는 평활 상수를 구하는데 시행착오를 거듭해야 하는 문제가 있으며, CensusⅡ는 직관에 기초한 방법이기 때문에 이론적인 근거에 의한 통계적 특성을 평가하기 어렵다[22]. 최근에는 계절적 시계열을 분석하는 방법으로 ARIMA 모형에 근거한 분활기법[18, 20], DDS를 이용한 방법[23]등이 제시되고 있으나, 전문적인 지식이 있는 사람이라도 여러 번의 시행착오 과정이 요구되어 일반인이 사용하기에는 어려움이 있다[11].

경제시계열 분석은 연간 자료에 의한 장기분석과 분기별 혹은 월별 자료에 의한 단기분석으로 구분된다. 전자의 경우에는 계절성의 문제는 전혀 발생하지 않는다. 따라서 원래의 시계열에 계절조정을 행하지 않고 기존의 시계열을 그대로 분석에 이용하더라도 분석에 지장을 받지 않는다. 그러나 후자의 경우는 계절요인이 추세순환 요인과 결합되어 나타나므로 기존의 시계열만으로 추세순환 요인을 명확하게 파악할 수 없다. 계절변동의 영향을 조정하지 않는 경우에는 계절적인 증감 경향을 그대로 경제의 실제추세로 간주하여 실제와는 전혀 다

른 판단을 함으로써 문제가 발생하게 된다. 따라서 단기적인 경제 분석과 예측을 위해서는 원래의 시계열로부터 계절요인을 제거하는 절차가 필수적이라 할 수 있으며 이를 계절변동조정이라 하고 계절요인이 제거된 시계열을 계절변동조정 시계열이라 한다. 오랫동안 계절변동조정의 필요성이 인식되어 왔으나 계절변동조정 방법에 대한 논의는 아직도 계속되고 있고 현재도 어느 방법이 다른 방법보다 뚜렷하게 통계적으로 우수하다고 판명된 것은 없으며 각각의 시계열 자료를 경험적 방법으로 분석하고 있는 실정이다. 따라서 본 연구에서는 A사의 전자제품으로 계절성 상품인 냉장고의 판매매출액 자료에 X11-ARIMA 방법을 적용하여 계절조정하고 이 방법에 의해 산출된 불규칙요인의 임의성과 계절요인을 통계적 방법으로 분석하여 계절조정결과를 평가하고자 한다. 또한 계절조정 결과를 향상시키기 위하여 X11-ARIMA 방법에 기초한 적합한 모형을 이용한 자기상관과 계절성을 고려한 수요예측방법을 제시하고자 한다.

2. 시계열의 이론적 배경과 분석모형

2.1. 시계열의 분석모형

시계열 분석이란 독립변수와 종속변수간에 성립되는 인과관계적 함수관계를 파악한 연후에 이에 근거하여 미래예측을 수행하는 회귀분석과는 달리 단지 분석대상이 되는 변수의 과거 형태 내지 관측값 만으로 미래예측을 실시하는 방법론이라 할 수 있다. 이러한 시계열분석은 경제이론 등에 근거하지 않으면서도 단기예측에 있어서 전통적 회귀분석보다도 높은 정확성을 갖는 것으로 알려져 있다. 초기단계의 시계열 분석은 외삽법이나 평활법 등에 의해 순환변동 또는 계절변동 등의 내용을 파악한 다음 미래예측을 수행하는 단계에 머물렀으나 이제는 Box-Jenkins에 의해 제시된 ARIMA 모형을 이용하여 불규칙 변동까지도 고려하는 보다 정밀한 방법론이 일반적으로 사용되고 있다. 모든 종류의 ARIMA 모형은 확률적 과정(Stochastic process)의해 생성 될 수 있는데, 이러한 확률과정 개념은 기본적으로 확률오차 또는 백색오차(white noise)라 불리는 확률과정 개념을 이용하여 정의된다. 예컨대, 가장 대표적인 두 가지 확률과정인 자기회귀 확률과정(AR process)과 이동평균 확률과정(MA process)은 각각 아래와 같다.

$$Y_t = C + \phi_1 Y_{t-1} + \varepsilon_t \ , \ Y_t = C + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

(단, ε_t 는 백색오차)

그리고 이러한 MA 또는 AR과정의 차수를 변경하거나 서로 결합시킴으로써 다양한 시계열 모형을 창출해 낼 수 있는 것이다[9].

2.2 주기적 변동요인과 계절 변동요인

많은 시계열들은 추세변동 이외에 주기적인 순환변동에 영향을 받는다. 이러한 순환변동중 대표적인 것은 1년 주기를 갖는 계절 변동이며, 모든 순환 변동 중 가장 짧은 주기를 갖는다. 이러한 순환변동에서와 같은 차원의 분석이 1930년대까지 주류를 이루어 왔다. 그러나 이러한 방법의 한계성과 문제점이 노출되었고, 이를 해소하기 위한 노력이 경주된 결과 근래에 이르러 Box-Jenkins의 ARIMA모형에 의한 분석방법이 정착하게 되었다. 계절 변동요인에 관한 분석은 궁극적으로 그 변동내용과 변동정도를 지수화하여 파악함과 동시에 이 지수를 이용하여 원 시계열에서 계절변동요인을 제거시키거나 계절변동이 제거된 시계열에 계절변동 내용을 포함시키는 것을 주 내용으로 하고 있다. 이러한 목적을 위해 도입된 분석방법은 최근 Box와 Jenkins에 의해 괄목할 만한 형태로 전개되었다. 즉, ARIMA 분석모형에다가 전통적인 계절조정지수 산출방법을 가미하여, 근대적인 계절변동의 분석작업이 시계열분석을 통하여 가능하게 되었다[9].

2.3. 추세 순환 조정방법

2.3.1 회귀분석 방법

추세요인을 나타내기 위해서는 일반적으로 다항식(polynomial)을 사용한다. 이 다항식은 임의의 함수 근사형이며, 원래의 함수는 이보다는 더 복잡한 형태일 것이다. 따라서 추정된 추세다항식의 계수들에 실제적 의미를 부여할 수 있는 경우는 매우 드물다.

시계열 $Z_1, Z_2 \cdots, Z_n$을 다음과 같이 p차 다항식과 오차 항으로 나타낼 수 있다고 하자.

$$Z_t = \beta_0 + \beta_1 t + \cdots + \beta_p t^p + a_t \tag{1}$$

단,{ a_t}는 오차항으로서 분산이 σ_a^2 인 백색잡음 과정이다. 이 모형을 사용해서 최소제곱 추정법이나 최우 추정법을 이용하여 추세 다항식을 구할 수 있다. 그러나 이 모형의 계수 추정량은 여러가지 문제점을 내포하고 있다. 첫째 : 계수의 추정 값들이 높은 상관관계를 가지고 있다. 둘째 : p차 다항식으로 추세 곡선을 추정한 뒤 다시 한 차수 높여 (p+1)차 다항식으로 추세 곡선을 추정하고자 할 때는 p차 추세 다항식을 전혀 이용하지 못하고, 처음부터 다시 계산해야 한다. 셋째 : 회귀식의 차수 p를 결정하는데 이 모형은 효과적이지 못하다. 이러한 이유로 인해 일반 다항식을 변형한 직교다항식을 사용해서 추세 다항식을 구한다. 각 열들이 서로 직교인 디자인 행렬을 찾으면 직교다항식(orthogonal polynomial)을 쉽게 구할 수 있다. 이것은 추세 다항회귀식의 독립변수들 $1, t, \cdots, t^p$을 아래의 조건을 만족하는 직교 독립변수들 $\phi_0(t)$, $\phi_1(t)$, $\cdots$, $\phi_p(t)$로 변환함으로써 가능하다

$$\sum_{t=1}^{n} \phi_i(t)\ \phi_k(t) \Rightarrow \begin{cases} = 0 & (i \neq k) \\ \neq 0 & (i = k) \end{cases} \tag{2}$$

p차의 추세 다항식을 직교 다항식을 사용하여 표기하면 다음과 같다.

$$Z_t = \gamma_0\phi_0(t) + \gamma_1\phi_1(t)\cdots + \gamma_p\phi_p(t) + a_t \tag{3}$$

여기에서 $\gamma_0, \gamma_1, \cdots, \gamma_p$는 계수 $\phi_0, \phi_1, \cdots, \phi_p$ 와 선형관계에 있는 새로운 계수이다. 이 직교 다항식을 이용한 회귀식의 디자인 행렬은 다음과 같다.

$$X = \begin{pmatrix} \phi_0(1) & \phi_1(1) & \cdots & \phi_p(1) \\ \phi_0(2) & \phi_1(2) & \cdots & \phi_p(2) \\ \vdots & \vdots & \cdots & \vdots \\ \phi_0(n) & \phi_1(n) & \cdots & \phi_p(n) \end{pmatrix} \tag{4}$$

직교성에 의하여 대각선 이외의 모든 항들은 0이므로 다음 식을 얻을 수 있다.

$$X'X = diag\ (a_{00}, a_{11}, \cdots, a_{pp})$$

$$단,\ a_{kk} = \sum_{t=1}^{n} \phi_k^2(t) \tag{5}$$

직교 다항식의 계수 벡터 $\underline{\gamma} = (\gamma_0, \gamma_1, \cdots, \gamma_p)^t$로 정의하면 $\underline{\gamma}$의 최소제곱 추정량은 다음과 같다.

$$\hat{\underline{\gamma}} = (X'X)^{-1}X'\underline{Z}$$

$$단,\ \underline{Z} = (Z_1, Z_2, \cdots, Z_n)^t \tag{6}$$

따라서 $\widehat{\gamma_k}$ 의 최소제곱 추정량은 다음과 같다.

$$\widehat{\gamma_k} = \frac{\sum_{t=1}^{n} Z_t\, \phi_k(t)}{\sum_{t=1}^{n} \phi_k^2(t)} \qquad k=1,2\cdots,\ p \tag{7}$$

$\hat{\underline{\gamma}}$ 의 공분산 행렬은 $(X'X)^{-1}\ \sigma^2$이므로 $\widehat{\gamma_k}$ 의 분산은 다음과 같다.

$$var(\hat{\gamma}) = \frac{\sigma^2}{a_{kk}} \tag{8}$$

또한 분산 σ^2 의 추정량은 다음과 같다

$$\hat{\sigma}^2 = \frac{1}{n-p-1}\left\{ \sum_{t=1}^{n} Z_t^2 - \sum_{k=0}^{p} a_{kk}\, \hat{\gamma}_{kk^2} \right\} \tag{9}$$

직교 다항식의 회귀 계수들이나 분산의 추정량을 간단한 형태로 구할 수 있는 것은 디자인 행렬의 직교성에 의해 가능하다. 일반적으로 차수 p를 사전에 알지 못한다. 이 경우 p차의 직교 다항식으로 추정한 추세 곡선이 잘 맞지 않아 (p+1)차로 시계열을 회귀 시키고자 할 때 단순히 $\gamma_0, \gamma_1, \cdots, \gamma_p$ 는 p차 회귀 다항식의 결과를 사용하고 (p+1)차 항의 계수는 다음과 같이 계산한다.

$$\hat{\gamma}_{p+1} = \frac{\sum_{t=1}^{n} Z_t \phi_{p+1}(t)}{\sum_{t=1}^{n} \phi_{p+1}{}^2(t)} \tag{10}$$

또한 직교다항식을 사용해서 구한 계수 추정량들 사이에는 상관관계가 없으므로 추정량을 해석하는데 용이하다. 회귀분석을 이용한 순환요인의 추정은 다음과 같다.
계열 $Y_0,\ Y_1, \cdots, Y_n$ 은 다음과 같은 모형을 갖는다고 가정하자

$$Y_t = f(t) + \mu_t \tag{11}$$

단, f(t) 는 순환변동 함수이고 { μ_t }는 분산 이 σ^2인 백색 잡음과정이다. 함수 f(t)는 주기가 알려져 있을 때 주기 함수로서 다음과 같이 나타낼 수 있다.

$$a_0 + \sum_{j=1}^{q} \{\alpha(k_j)\cos(2\pi k_j t/n) + \beta(k_j)\sin(2\pi k_j t/n)\} \tag{12}$$

여기서 $\{k_j/n : j=1,2,\cdots,q\}$는 f(t)의 주어진 주기들의 집합이며, k_j는
$I=\{1,\ 2,\cdots,[n/2]\}$ 의 원소이다. 이 집합 I 에 들어가지 않는 k_j에 대한 주기는 다루지 않는다. 그 이유는 이러한 값들을 사용하면 삼각함수의 직교성이 성립되지 않아서 각 주기에 대한 회귀 계수들의 추정량들이 상관관계를 갖기 때문이다.

2.3.2 이동평균 방법

이동평균 방법은 한 시점의 추세요인이나 계절요인을 추정할 때 회귀분석방법에서와 같이 모든 자료를 이용하지 않고 그 주위의 자료만을 이용한다. 시계열 $Y_1,\ Y_2, \cdots, Y_n$ 이 다음과 같은 모형을 갖는다고 하자.

$$Y_t = f(t) + \mu_t \tag{13}$$

단, f(t) 는 결정함수이고 μ_t는 분산이 σ^2인 백색 잡음과정이다. 이때 t시점의 추세요인 추정값 Y_t^* 은 다음과 같다.

$$Y_t^* = \sum_{s=-m}^{m} C_s\ Y_{t+s} \tag{14}$$

$$\text{단, } \sum_{s=-m}^{m} C_s = 1$$

이동평균 방법을 사용하면 원래의 시계열보다 분산은 줄어드나 편의는 커지게 된다.

$$E(Y_t^*) = E\left\{\sum_{s=-m}^{m} C_s Y_{t+s}\right\} = \sum_{s=-m}^{m} C_s\, f(t+s) \neq E\,(Y_t) = f(t) \tag{15}$$

$$Var\,(Y_t^*) = \sigma^2 \sum_{s=-m}^{m} C_s^2 < Var(Y_t) \tag{16}$$

이동평균 방법은 원래의 시계열에 계절요인이 어떠한 형태로 결합되어 있는가 또는 계절요인의 특성을 어떻게 보느냐에 따라 몇 가지 유형으로 나누어진다. 첫째, 시계열을 구성하고 있는 각 요인이 어떠한 형태로 결합되어 있는가에 가법형과 승법형으로 크게 나눌 수 있다. 일반적으로 승법형이 널리 사용되고 있으나 원래의 시계열에 '0' 이나 음수가 포함되어 있는 경우에는 승법형의 적용이 불가능하므로 가법형을 사용하고 있다. 둘째, 계절 패턴을 고정적인 것으로 보느냐 가변적인 것으로 보느냐에 따라 고정 계절형과 가변 계절형의 두 가지 조정 방법으로 나눌 수 있다. 전자는 계절요인이 상당 기간에 걸쳐 일정한 패턴을 유지하고 있다는 전제하에 이를 찾아내어 원래의 시계열을 조정하는 방법으로서 고정 계절값 방법, 고정 계절지수 방법 등이 이에 해당한다. 반면에 계절요인은 사회적 습관이나 전통, 제도 등 인위적인 요인과 기후변화 등 자연적 요인이 복합되어 나타나는 것이 일반적이기 때문에 계절요인의 패턴은 고정된 것이라기보다는 시간에 따라 변화 한다고 보는 것이 가변 계절조정 방법이다. 따라서 고정 계절지수가 실제 시계열의 움직임을 충분히 반영하지 못하는 문제를 해결하기 위해서 고안된 것이 가변형 계절지수 모형에 의한 각종 계절 변동조정 방법이다. 고정 계절조정 방법이 비교적 일찍이 개발된 방법임에 비하여 가변형 계절조정 방법은 많은 계산량을 필요로 하기 때문에 컴퓨터가 널리 보급된 최근에 발전하고 있으며, 그 대표적인 것이 미국 상공부의 센서스 국법이다. 이동평균 방법을 네 단계로 나누어 보면 다음과 같다.

단계1.추세요인, 순환요인, 계절요인 및 불규칙요인으로 구성된 원래의 시계열에 이동평균 방법을 적용하여 계절요인 및 불규칙을 제거하여 추세순환 요인을 얻는다.

단계2.원래의 시계열로부터 앞에서 얻은 추세순환 요인을 제거하여 계절 및 불규칙 요인을 얻는다.

단계3.위에서 산출한 계절 및 불규칙 요인을 매년 동일한 분기별 (월별 시계열에 대하여는 월별)로 이동평균하여 불규칙 요인을 제거하여 계절 요인만을 얻는다.

단계4.원래의 시계열로부터 위에서 추정한 계절요인을 제거함으로써 계절변동 조정 시계열을 얻는다.

오늘날 사용하는 대부분의 계절변동조정 방법들은 이동평균 방법의 범주에 속하며 널리 사용되고 있는 것은 X11 방법과 X11-ARIMA 방법이다. 이동평균방법을 이용할 때, 보다 양호한 계절변동조정 시계열을 얻기 위해서는 각종 이동평균 방법의 서로 다른 성격을 충분히 이해하는 것이 중요하다. 그런데 여러 이동평균 방법은 각각의 장단점이 있고 서로 상충 관계에 있기 때문에 어떤 방법이 절대적으로 우월하다는 기준은 있을 수 없다. 결국 개별 시계열의 특성에 비추어 어떠한 형태의 이동평균 방법이 보다 합리적인 결과를 가져올 것인가를 세심하게 검토하여야 한다. X11 방법과 X11-ARIMA 방법에서 사용되는 이동평균 방법은 가중값의 형태에 따라 대칭형과 비대칭형으로 분류될 수 있다. 대칭형 이동평균 방법은 (2n +1)개항을 이동평균 할 경우 중간항을 제외한 처음과 마지막의 n개항이 대칭적으로 동일한 가중 값을 부여받는 방법이며, 비대칭형 이동평균 방법은 이동평균 과정에서 처음 또

는 마지막의 n개항이 손실됨으로써 처음 또는 마지막 항에 가까운 일부 항에 큰 가중값이 부여되는 방법을 말한다. 그런데 두 가지 이동평균 방법의 가중값의 합계는 모두 1 이기 때문에 어떤 방법으로 원래의 시계열을 이동평균 하여도 그 결과의 평균값은 변하지 않는다. 다만, 이동평균 방법에서 유의해야 할 점은 원래의 시계열과 이동평균된 시계열을 비교할 때 시간상의 국면이동(phase shift)이 일어나지 않도록 해야한다. 즉, 시계열의 상승과 하락이 시간상으로 뒤바꾸지 않아야 한다.

2.4 계절변동 조정방법

현재 대부분의 계절변동조정 방법은 단변량 시계열모형(univariate time series model)에 기초를 두고 있는데 단변량 시계열 모형에서 시계열은 시간의 함수인 규칙적 부분(systematic part) 과 확률 법칙에 따르는 임의적 부분(random part)으로 구성되어 있다. 임의적 부분은 일정한 분산과 평균0, 비자기상관 등의 특징을 가지며 시계열의 전기간에 걸쳐 동일한 분포를 이룬다고 가정된다. 규칙적 부분은 여러 요인에 의해 구성되어 있는데 일반적으로 추세요인, 순환요인, 계절요인으로 구분하고 있다. 추세요인은 시계열이 장기간에 걸쳐 증가하거나 감소하는 현상을 말하며 순환요인은 다양한 진폭과 주기로 증가와 감소를 반복하는 현상을 말한다. 일반적으로 순환요인의 주기는 길기 때문에 추세요인의 일부로서 취급되는 경우가 많다. 계절요인은 시간별 전력소비량, 월별 통화량, 분기별 수출액 등과같이 일정한 주기를 가지고 반복하는 현상을 말하는데 순환요인과의 차이점은 그 주기가 짧다는 것이다[2]. 계절조정모형에서 시계열 구성요인의 추정방식은 크게 회귀분석 방법과 이동평균 방법의 두 가지로 나누어진다. 회귀분석 방법은 시계열의 구성 요소중 규칙적 부분인 추세요인, 순환요인 및 계절요인들이 시계열의 전기간에 걸쳐 결정함수(deterministic function)에 의해 추정이 가능하다고 가정하고 있다. 반면에 이동평균방법은 시계열의 구성요인들은 시간의 평활함수(smoothing function)로서 이 요인들을 시계열의 전기간에 걸친 단순 함수로 가정하는 것이 아니라 시점에 따라 변화한다고 가정하는 것이다.

2.4.1 계절 계열과정(seasonal process)에 대한 이론적 acf와 pacf

(1) 비 계절적 1차 자기회귀 계열과정 (non-seasonal AR(1)

전통적인 시계열 분석에서 계절 지수를 산정하는 기본적인 방법론은 계절변동이 1년을 주기로 규칙적으로 반복되는 성격을 가진 것으로보고 1년간의 시차를 가진 관측치를 보통 5개년간 평균하여 계절 지수로 산출하는 것이었다[1]. 이러한 내용을 ARIMA모형 분석에 도입하기 위하여 계절변동이 있는 시계열의 acf와 pacf를 살펴 보는것도 중요한 의미를 갖는다. 비계절적 1차 자기회귀 계열과정(non-seasonal AR(1))은 ACF가 지수적으로 감소하여 0에 접근하는 것을 말하며, 안정적 자기회귀 계절계열과정(stationary seasonal AR(1) process)은, 예컨대 분기별 자료를 사용할 경우 이론적인 ACF는 점차 감소하여 0에 접근한다는 점에서는 일반 AR(1)과 같으나 계절 시차(seasonal lags)가 s = 4, 8. 12 … 인 경우에 한해서만

지수적으로 감소 한다는 점에서 차이가 있다. 두 과정의 특성상의 차이는 <표 1>에서 찾아 볼수 있다.

<표 1> 계절적, 비계절적 AR(1)의 이론적 ACF (ϕ_1=0.7로 가정) [9]

k	계절적 AR의 ρ	비계절적 AR의 ρ
1	0.7	0
2	0.49	0
3	0.34	0
4	·	0.7
5	·	0
6	·	0
7	·	0
8	·	0.49
9	·	0
10	·	0
11	·	0
12	·	0.34
13	·	0

(2) 계절적 S차 자기회귀 계절계열과정

일반적으로 “s ”의 시차를 갖는 자기회귀 계수로 표현되는 계절 계열과정(seasonal process)을 정의하면 다음과 같다.

$$Y_t = \delta + \Phi_s Y_{t-s} + \varepsilon_t \tag{17}$$

$$(1 - \Phi_s B^s)\ y_t = \varepsilon_t \tag{18}$$

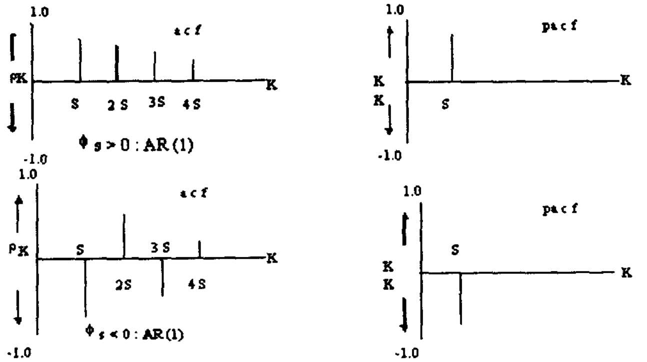
단, $y_t = Y_t - E[Y_t]$

한편, 시차“s”의 MA과정으로 표현되는 계절계열과정은 아래와 같이 표현이 가능하다.

$$Y_t = \delta - \Theta_s \varepsilon_{t-s} + \varepsilon_t \tag{19}$$

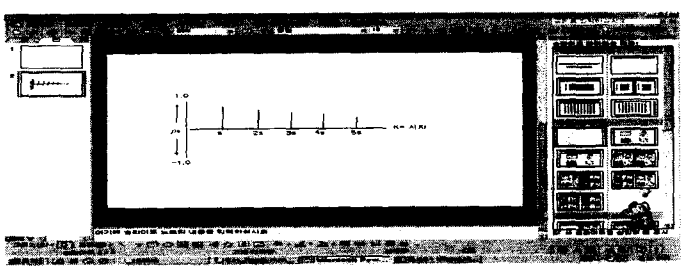
$$y_t = (1 - \Theta_s B^s)\ \varepsilon_t \tag{20}$$

이제 AR(1)과 MA(1)의 경우 Φ_s 와 Θ_s 부호의 다양한 가정하에 이론적 ACF와 PACF를 도표화하면 <그림 1>와 같다[9].



<그림 1> 안정적 계절 AR(1)과 MA(1)의 이론적 ACF와 PACF

위<그림 1>에서 비 계절적 계열(nonseasonal) AR(1)과 MA(1)의 유사성과 차이점은 분명하다. 그러나 실제에 있어서 추정된 ACF와 PACF를 가지고 계절 계열모형(seasonal model)을 식별하는 것은 훨씬 어렵다. 계절적 변동을 지니는 많은 실현값들이 아울러 비 계절적 계열 형태를 동시에 포함하고 있기 때문이다. 즉, 추정된 ACF와 PACF는 비계절적 계열과 계절적 계열 두 요인을 모두 반영하고 있다. 이들로부터 시각적으로 계절적 계열의 PACF와 비 계절적 계열의 PACF를 구분하는 것은 매우 어렵다. 한편, 불안정적 평균을 지니는 계절적 계열과정은 불안정적 비 계절적 계열과정(nonstationary nonseasonal process)과 역시 형태상 유사한 ACF를 갖는다. 즉, 불안정적 평균을 지니는 계절적 계열과정은 역시 ACF가 "0"으로 빨리 이동하지 않는 특징을 갖게 된다. 어느 실현값에서 추정된 ACF가 애래<그림 2>의 현상을 보일 경우 계절별 차분이 요구된다[9].



<그림 2> 불안정적 평균을 지니는 계절적 계열과정의 ACF

2.5. X11-ARIMA 방법

X11 방법의 취약점을 보완하기 위하여 카나다 통계청(Statistic Canada)에서 단기예측을 위한 시계열을 분석 기법인 [Box-Jenkins]의 ARIMA모형(autoregressive integrated moving average model)과 X11 방법을 결합하여 X11-ARIMA 방법을 개발하였다[13, 15,

16]. X11-ARIMA 방법은 ARIMA모형에 의해 최소 평균제곱 오차(minimum mean square error)를 가지는 예측값을 산출, 원래의 시계열을 연장하여, 연장된 시계열에 X11 방법을 적용하므로 계절요인의 수정 폭을 극소화시키는 이점이 있다. 이 방법은 실효성이 인정되어 1975년 캐나다 통계청에서 공식적으로 채택되었고 그 밖의 여러 나라에서도 가장 발전된 시계열의 계절변동 조정방법으로서 널리 사용되고 있다.

2.5.1 ARIMA모형의 식별

X11 방법은 추세순환 요인을 산출하기 위하여 중심화 이동평균을 수행하는 과정에서 시계열 양끝의 자료가 손실되므로 최근 계절요인 추정값의 신뢰도가 저하된다. X11-ARIMA 방법은 이 단점을 극복하기 위하여 ARIMA 모형을 사용해서 시계열의 양끝을 각각 1년씩 연장한 후 X11방법을 적용한다. X11-ARIMA 방법은 ARIMA 모형에 의해서 시계열의 양끝이 각각 연장되어 중심화 이동평균 과정에서 결측 자료가 발생하지 않기 때문에 X11 방법에 비하여 보다 진일보한 계절변동조정 방법이다. 이 때 과거 1년간을 연장하는 것은 5~6년 사이의 시계열에만 적용되는데 이는 과거 연장값이 시계열에 새로운 관찰값이 추가될 때마다 변하고 또한 현재의 계절조정 결과에도 영향을 미치기 때문이다. Dagum[19]은 X11방법의 약점을 보완하기 위하여 ARIMA모형을 예측 수단으로 채택한 이유를 다음과 같이 들고 있다.

① ARIMA 모형은 다른 설명변수를 도입하는 대신에 변수의 과거 관찰값과 오차항만을 가지고 시계열에 적합한 모형을 설정 할 수 있는 간편한 방법이다.

② ARIMA 모형은 1년 혹은 2년간의 자료가 새로이 추가되어도 이미 선정된 모형의 모수추정 값이 크게 변화하지 않기 때문에 계절변동조정 시계열의 안정성이 유지될 수 있다.

③ ARIMA 모형은 시계열의 특성을 적은 수의 모수로 표현 할 수 있다.

④ ARIMA 모형으로부터 생성된 예측값은 최소평균제곱오차 개념에 의해서 도출된다.

⑤ ARIMA 모형에서 생성된 예측값은 시계열의 움직임을 합리적으로 반영한다. 즉 예측값이 실제값과 다르다고 하더라도 분기별 또는 월별 변동패턴을 합리적으로 반영한다.

ARIMA 모형은 개별 시계열의 특성에 따라 달라지게 되므로 각각의 시계열에 적합한 ARIMA 모형을 식별해 주어야한다. X11-ARIMA 방법에서는 이러한 점을 고려하여 적용범위가 큰 모형들을 내장하고 있어 적합한 모형을 자동적으로 선택할 수 있고 또한 임의로 개별 시계열의 특성에 알맞은 ARIMA모형을 식별하여 계절변동 조정을 할 수 있다. ARIMA모형의 자동 선정을 위하여 시계열이 최소한 5년 이상의 시계열이어야 하고 , 15년 이상인 시계열에 대하여는 ARIMA모형을 식별 할 때는 최근의 15년만이 이용된다. 내장시킬 ARIMA모형을 선정하기 위하여 몇번의 실증분석이 수행되었다. Lothian and Morry[21]은 174개 카나다 경제 시계열의 15년간의 월별 및 분기별 자료를 대상으로 실증 분석을하여 12개의 계절 ARIMA모형 중에서 가장 적합성이 높은 것으로 판단되는 3개의 모형을 선택하였다.

2.5.2 X11 방법

X11방법에서는 시계열을 구성하는 요인들을 추세순환요인, 계절요인, 불규칙요인 및 요일구성요인(trading day variation factory)으로 분류한다. 또한 각 요인들과 관찰과의 관계에 대해 다음과 같은 승법형 모형 또는 가법형 모형을 가정하고 있다.

$$O_t = C_t\, S_t\, I_t \tag{21}$$

$$O_t = C_t + S_t + I_t \tag{22}$$

단 O_t : 원래의 시계열　C_t : 추세순환요인

S_t : 계절요인　I_t : 불규칙요인

이 2개의 모형 중 선택된 하나의 모형에 대해 X11 방법은 다음과 같은 이동평균 과정과 이상값(outlier) 수정과정으로 13단계의 계산 과정을 2회 반복한다. 13단계의 계산과정을 월별자료에 승법형 모형을 적용하여 설명하면 다음과 같다.

단계1. 원래의 시계열을 중심화 12개항 이동평균하여 잠정적인 추세순환 요인을 산출한 후 원래의 시계열을 나누어 계절 및 불규칙요인을 구한다.

단계2. 제1단계에서 구한 각 월의 계절 및 불규칙을 가중 5 개항 이동 평균하여 계절요인의 1차추정 값을 산출한다. 이때 이동평균 과정에서 손실되는 양끝 2개항은 가장 근접한 항으로 대체된다.

단계3. 제2단계에서 구한 계절 요인을 중심화 12개항 이동평균 한다. 이때 12개항 이동평균 과정에서 시계열의 양끝에 6개월의 결측 자료가 발생하므로 이를 보충하기 위하여 처음 [마지막]의 결측자료는 바로 다음 해 [전년도] 동월의 값으로 대체된다. 계절 요인을 이 이동평균 값으로 나누어 계절요인의 2차추정 값을 구하고 일년간의 계절 요인들의 합계가 12가 되도록 조정한다.

단계4. 제1단계에서 구한 계절 및 불규칙요인을 제3단계에서 구한 계절요인을 나누어 불규칙요인을 산출한다.

단계5, 제4단계에서 산출한 불규칙요인의 이상 값을 다음과 같이 조정한다. 산출된 불규칙요인에 대하여 5년간 계열의 표준편차 σ를 계산하고 이 5년간의 가운데 해의 불규칙요인이 2.5 σ 를 넘으면 이를 제거한 후 표준편차 σ를 다시 계산한다. 다시 계산된 표준편차 σ를 기준으로 2.5 σ 이상인 불규칙요인에는 가중값을 0을, 1.5 σ이하인 불규칙요인에는 가중 값을 1을, 1.5 σ와 2.5사이에 있는 불규칙요인에 대하여는 0과 1사이의 가중 값을 부여한다.

단계6. 제5단계에서 1보다 작은 가중값이 부여된 항의 계절 및 불규칙 요인은 해당 월의 계절 및 불규칙요인에 가중 값을 곱한 값과 가중값 1을 부여받은 전후 인접한 각 2개항과의 평균값으로서 대체한다. 이때 처음 2개년에 3차년의 표준편차를, 마지막 2개년에는 끝에서 3번째 년도의 표준편차를 이용한다. 처음과 끝의 2개년에 있는 각 월별 이상 값을 대체할 계절 및 불규칙요인에 가중 값을 곱한 값과 가중값 1을 부여받은 인접한 4개항과의 평균을 구한 후 이 값으로 대체한다.

단계7. 이상 값을 대체한 계절 및 불규칙 요인을 월별로 가중 5개항 이동평균하여 잠정적인 계절요인을 산출한다. 이 때 손실된 양끝 2개항이 보충을 위해서 제2단계와 같은 방법을 사

용한다.

단계8. 계절요인에 대해서 제3단계의 계산과정을 반복한다.

단계9. 잠정적인 계절변동조정 시계열을 구하기 위하여 원래의 시계열을 잠정 계절요인으로 나눈다.

단계10. 잠정 계절변동조정 시계열에 9개항, 13개항 또는 23개항 헨더슨 이동평균 방법을 적용하여 추세순환요인을 산출하고 이것으로 원래의 시계열을 나누어 계절 및 불규칙요인의 추정 값을 구한다.

단계11. 계절요인의 추정 값을 얻기 위하여 각 월별 계절 및 불규칙요인을 가중 7개항 이동평균 한다.

단계12. 제3단계로 되돌아가서 같은 과정을 반복한다.

단계13. 원래의 시계열을 계절요인으로 나누어 계절변동조정 시계열을 산출한다.

X11 방법은 기본적으로 각 종 이동평균 방법을 수차례 반복함으로써 4개의 시계열 구성요인을 분해하고 동시에 불규칙요인의 이상값을 제거하여 보다 안정적인 계절변동조정 시계열을 얻고자 하는 방법이다. 그러나 이 방법은 다음과 같은 단점을 가지고 있다.

첫째, 원래의 시계열의 분해에 관한 명확한 모형(explicit model) 이 없다. 둘째, 여러 번의 이동평균 과정에서 좌우 대칭형 가중 값을 부여받는 시계열 중앙부분 관찰값 과는 달리 양 끝의 관찰값은 좌우 비대칭형 가중 값을 부여받으므로 중앙부분 관찰값에 비해 각 구성요인의 추정값에 대한 신뢰도가 떨어지고 아울러 현재의 시계열에 새로운 자료가 추가되면 계절요인의 수정 폭이 크고 그 수정에 의해 계절변동조정 시계열의 움직임이 달라지는 경우도 있다. 이와 같이 이동평균 방법이 가지는 두 번째 단점을 보완하기 위하여 1974년 Dagum[19]이 X11-ARIMA라는 방법을 개발하였다.

2.6. X11-ARIMA 방법에 의한 계절조정과정

X11-ARIMA 방법의 계산 과정은 다음과 같이 7개의 부분으로 대별된다.

A. 사전조정 요인에 의한 조정

B. 잠정 요일구성요인, 이상값 수정을 위한 가중값 등의 계산

C. 최종 요일구성요인, 이상값 수정을 위한 가중값 등의 계산

D. 최종계절요인, 추세순환요인, 불규칙요인 및 계절변동조정 시계열의 계산

E. 이상 값을 수정한 원래의 시계열, 최종 계절변동 조정 시계열의 계산

F. 추세순환요인 지배기간 (month for cyclical dominance : MCD) 에 의한 이동 평균 및 각 요인의 특성값 등의 계산

G. 도표의 작성

7개의 부분 중에서 A부분은 각 시계열에 대해 사전 지식이 있을 경우 이를 처리하는 과정이고 B-D부분은 계절변동 조정 시계열을 산출하기 위한 중심적인 계산과정이다. 위의 계산과정은 X11방법과 X11-ARIMA방법이 동일하나 X11방법은 주어진 원래의 시계열만을 이용하여 계절변동 조정 시계열을 산출하는데 비하여 X11-ARIMA 방법에서는 원래의 시계열에

적합한 ARIMA모형에 의하여 시계열을 연장하여 계절 조정한다는 점이 다르다.

2.6.1 사전조정요인

원래의 시계열에서 계절요인을 산출하기에 앞서 원래의 시계열의 변동에 관해 사전지식을 가지고 있을 경우 이를 먼저 처리할 수 있다. 원래의 시계열 변동에 관한 사전 지식이란 원래의 시계열이 특정기간 중 어떤 요인에 의해 심한 불규칙성을 나타내고 있을 경우에 그 정도를 측정할 수 있거나 또는 요일별 가중 값을 사전에 추정할 수 있는 경우를 말하는 것으로 이러한 불규칙요인을 원래의 시계열로부터 사전에 조정할 수 있다. 우리 나라의 경우 추석 또는 구정이 들어 있는 월이 어느 월인가에 따라 발생하는 시계열의 변동이 여기에 해당된다. 사전에 파악이 가능한 요인을 영구적 사전 조정 요인과 일시적 사전 조정요인으로 구분하고 있는데 영구적 사전 조정요인이란 어떤 불규칙요인을 원래의 시계열에서 영구히 제거하고 그 이후 계절변동 조정 시계열에 환원시키지 않는 조정요인을 말하며, 일시적 사전 조정요인이란 원래의 시계열에서 제거하되 최종 계절변동 조정 시계열을 산출하였을 때 그 영향을 다시 환원시키는 조정 요인을 말한다.

2.6.2 시계열 구성 성분의 분해

X11-ARIMA 방법은 계절요인의 산출을 일차적으로 목표로 하고 있으나 이를 추출하기 위한 계산과정에서 원래의 시계열을 각 요인별로 분해하기 때문에 사실상 시계열 분해방법이다. 일반적으로 시계열 구성요인은 추세요인, 순환요인, 계절요인 및 불규칙요인의 4가지 요인으로 구성되어 있다고 간주하고 있다. 그러나 X11-ARIMA 방법에서는 일반 분해 방법과 비교할 때 약간의 차이점을 지니고 있다. 즉 추세와 순환요인을 하나로 묶어 추세순환요인으로 하는 반면 불규칙요인 중 요일 구성에 따른 변동요인을 순수한 불규칙요인과 별개로 구분하여 추출한다. 추세와 순환요인을 구분하지 않는 것은 순환요인이 추세의 일부 특성과 동일할 뿐만 아니라 그 진폭이나 주기의 상대적인 크기를 정확하게 분리해 내기 어렵기 때문에 이를 따로 분리하기 위한 과정에서 오히려 다른 요인들의 정확한 추출을 어렵게 만들 우려가 있다는 점에 근거를 두고 있다. 요일 구성요인이란 조업일수에 의한 변동요인을 말한다. 즉 특정월의 일수는 일정하나 매월 시작되는 요일은 해마다 변동하게 되어 조업일수가 달라지게 된다. 이 변동요인을 불규칙으로부터 추출해내는 것이 요일 구성요인 조정이다. 요일 구성요인은 두 부분으로 구별하여 조정한다. 즉 요일구성에 의한 변동요인을 사전에 어느 정도 산출해 낼 수 있다고 가정하는 사전 요일구성요인 조정과 불규칙요인에 포함되어 있다고 가정하는 잔존 요일구성요인 조정으로 나누어진다. 승법형 모형의 경우 사전적으로 각 요일의 가중값이, 합이 7이 되도록 주어졌다면 사전 요일구성요인은 다음 식에 의하여 계산된다.

$$M_i = \frac{X_{1i}D_{p1} + X_{2i}D_{p2} + \cdots + X_{7i}D_{p7}}{N_i} \tag{23}$$

단, M_i : i월의 사전 요일구성요인

$X_{j\,i}$: i월의 j요일 수

$D_{p\,j}$: j요일의 사전가중 값

N_i : i월의 일수

이때 사전 요일 구성요인의 조정은 원래의 시계열을 사전 요일구성요인으로 나누어줌으로써 이루어진다.

2.6.3 이상값의 수정

수 개년 이상의 경제시계열을 이용하여 고정 계절변동요인 산출방법에서는 각 월의 자료에 포함되어있는 불규칙요인이 계절요인을 산출하는 과정에서 각 월별로 평균함에 따라 자연히 제거된다. 그러나 X11-ARIMA방법은 특정 년, 특정 월의 계절요인이 별도로 계산되어 나오므로 불규칙요인의 제거가 문제된다. 따라서 원래의 시계열 구성성분인 불규칙 요인 중 이상 기후나 파업 등에 의해 발생된 불규칙요인을 제거하기 위한 계산과정이 필요한데 이 계산과정을 이상 값(outlier)의 수정 과정이라 한다. 여러 차례 반복되는 계산 처리과정에서 분리되어 나오는 각 월의 불규칙 요인을 5개년의 구분하여 표준편차를 계산하고, 이를 기준으로 일정한 한계를 설정하여 불규칙요인이 이 관리한계를 넘어서는 경우에는 이를 이상값으로 보아 해당하는 월의 불규칙요인을 가중 평균된 인접 년의 불규칙 요인으로 대체한다. 이렇게 하여 최종적인 계절요인을 산출할 때 원래의 시계열 구성 성분 중 불규칙요인을 크게 내포하고 있는 이상 값을 제거하여 순수한 계절요인만을 산출한다.

2.7. X11-ARIMA 계산과정에 있는 통계적 검정방법

X11-ARIMA 방법은 보다 정확한 계절요인을 추정하기 이해 다음과 같은 다양한 통계적 검정 수단을 이용한다.

2.7.1 ARIMA 모형선택에 관한 통계적 검정

(1) 절대 평균예측 오차

X11-ARIMA 방법에 의하여 선택한 모형으로부터 최근 3년간 추정값에 대해 절대 평균예측 오차를 구하여 이값이 15%보다 크면 선택된 ARIMA모형이 부적합한 것으로 판단한다.

(2) 모형의 유의검정

ARIMA 모형에 의한 잔차들이 임의적인가를 퍼트멘토우 통계량에 의해 유의수준 5% 하

에서 검정한다.

(3) 과잉차분 검정

과잉차분이란 ARIMA (p. d. q) > (P. D. Q)s 모형에서 추정된 일반 이동평균 모수 혹은 계절 이동평균 모수의 합이 0.9보다 큰 경우를 말하는데, 선택된 모형이 과잉차분 되어 있으면 보다 더 간단한 모형으로 변형 될 수 있기 때문에 부적합한 모형으로 판단한다.

2.7.2 계절성 존재에 관한 검정

X11-ARIMA 방법은 보다 정확한 계절요인을 추정하는 한편 개별 시계열의 계절성을 통계적으로 식별해 낼 수 있는지를 판단하기 위하여 다음과 같은 다양한 통계적 검정수단을 이용한다.

(1) 안정적 계절성에 대한 F검정

안정적 계절성이란 시계열에 계절성이 주기적으로 존재하는 것을 말하는데 계절성이 불안정적일때는 계절요인의 식별이 불가능해진다. X11-ARIMA 방법에서는 최초로 산출된 미수정 계절 및 불규칙요인(B1)과 최종적으로 산출된 미수정 계절 및 불규칙요인 (D8)에 대하여 다음과 같이 모형을 설정하고 일원 분산분석에 의한 F검정을 유의수준1% 하에서 실시하여 안정적 계절성의 유무를 판단한다.

$$SI_{ij} = SI_j + I_{ij} \tag{24}$$

단, SI_{ij}: i년 j월의 계절 및 불규칙요인

SI_j : j월의 계절 및 불규칙 요인의 평균

I_{ij}: i년 j월의 불규칙요인

$I_{ij} \sim N(0.\sigma^2)$

$Cov(I_{ij},\ I_{k1}) = 0 \quad ij \neq k1$

(2) Kruskal -Wallis 의 χ^2 검정

일원분산 분석을 할 때 각 표본들은 동일한 분산을 가진 모집단으로부터 추출된 것이라는 가정이 필요하기 때문에 안정적 계절성에 대한 검정을 할 때 비모수적 방법인 Kruskal-Wallis 검정도 함께 수행한다.

(3) 불규칙 요인의 계절성에 대한 검정

D11 즉 최종 계절변동조정 시계열에서 추세순환 요인을 제거한 후 불규칙 요인속에 잔여 계절성이 존재하는지 여부에 대한 F검정을 수행한다. 추세순환 요인을 제거하는 한 방법으로 월별 자료인 경우에는 3개월 시차로 1차 차분하고, 분기별 자료인 경우에는 1분기 시차로 차분 한다. 이 결과 산출된 시계열에 대하여 시계열 전기간 및 최근 3개년 시계열에 대하여 F 검정을 한다.

3. 모형적용과 계절변동조정

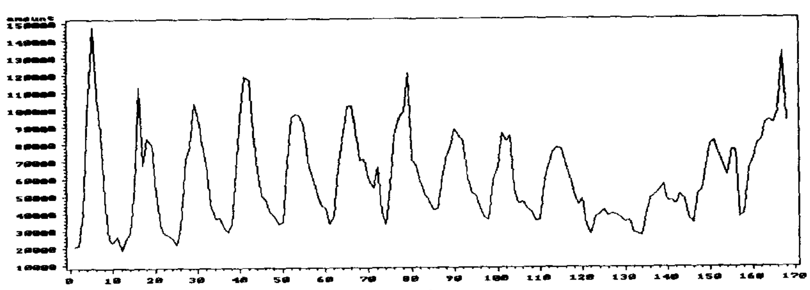
3.1 분석 개요

ARIMA 모형은 개별 시계열의 특성에 따라 달라지게 되므로 각각의 시계열에 적합한 ARIMA 모형을 식별해 주어야 한다. X11-ARIMA 방법에서는 이러한 점을 고려하여 적용범위가 큰 모형들을 내장하고 있어 적합한 모형을 자동적으로 선택할 수 있고 또한 임의로 개별 시계열의 특성에 알맞은 ARIMA 모형을 식별하여 계절변동조정을 할 수 있다. ARIMA 모형의 자동선정을 위하여 시계열이 최소한 5년 이상의 시계열이어야 하고, 15년 이상인 시계열에 대하여 ARIMA모형을 식별할 때는 최근의 15년만이 이용된다. 본 연구에서 제시된 계절변동조정 시계열의 실증분석에 이용된 <표 2>의 A사의 전자 제품인 계절상품 냉장고의 판매매출액 자료는 1988년 1월부터 2001년 12월까지의 자료이며, 2002년도의 실적 자료는 집행되지 않아 14년간만의 자료를 적용하였다. 조업일수와 밀접한 관련이 있는 요일구성요인은 본 연구에서 제시한 계절상품 판매매출액을 이용한 계절변동 조정 시계열에서는 제외하였다.

3.2 계절변동조정 결과분석

3.2.1 데이터 기술

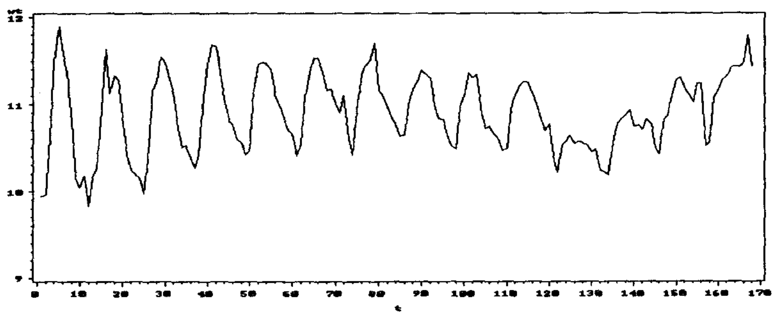
<그림 3>은 1988년 1월부터 2001년 12월까지의 14년(168개월)간의 A사 계절상품 냉장고의 판매매출액에 대한 관측 자료를 시계열 분석하여 시간으로 그래프로 표현한 것이다. 아래 <그림 4>을 살펴보면 계절성 상품인 냉장고라는 특수성이 계절적인 영향을 받고 있음을 알 수 있다. 88년 올림픽을 맞이하여 경기가 살아나서 전반적인 판매량이 평균에 비해 높게 나타나고 있으며, 98년에 들어서는 특정 변수인 IMF를 맞이하여 판매량이 격감한 것으로 보여진다. 또한 최근에 들어서 가파른 상승 추세를 나타내고 있는 것으로 보아 98년 IMF이전의 수준으로 회복 단계에 접어들어 판매매출액이 급격히 증대하고 있으며, 1년을 주기로 일정한 계절적 패턴을 가지고 있음을 알 수 있다.



<그림 3> A사 판매매출액 원시계열 불규칙요인 변동추이

3.2.2 모형식별

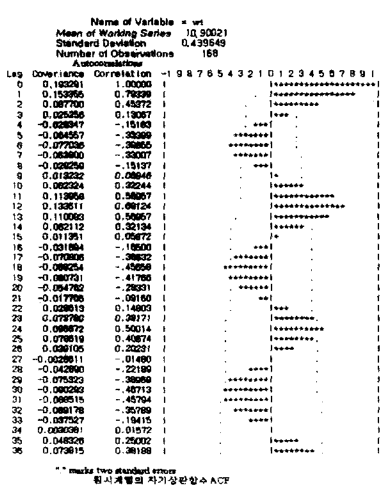
X11-ARIMA 모형을 SAS를 이용하여, 적합하기 위하여는 ARIMA모형을 먼저 적용한 후 X11-PROC를 사용하여 시계열의 성분을 추세순환 요인과 계절적요인, 불규칙 요인으로 구분한 데이터 값을 새로 얻을 수 있다. 먼저 적절한 ARIMA 모형을 적합하기 위해서는 자료의 분산 안정화가 필요하다. 따라서 <그림 4> 에서와 같이 모든 자료에 대수변환(log)을 취하여 분산을 어느 정도 평활화 하여 일정하게 변환하였다.

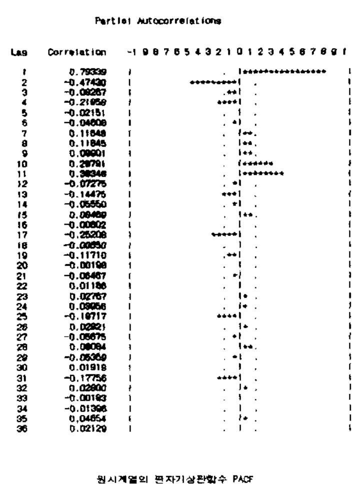


<그림 4> 원시계열 대수변환그래프

3.2.3 모형적합검증

적절한 ARIMA모형을 적합하기 위하여 <그림 5> 원시계열의 자기상관함수 ACF와 <그림 6> 편자기상관함수 PACF를 SAS를 이용하여 구한 결과 아래와 같다.





<그림 5>원시계열 자기상관함수 ACF <그림 6>원시계열 편자기상관함수 PACF

위에 주어진 <그림 5>의 원시계열의 자기상관 함수 ACF가 "+"부호부터 시작하되 부호 를 바꾸어 가며 계절 시차들인 6. 12, 18의 배수에 해당 되는 시차에서 점차 0에 수렴해 갈 뿐, 스파이크가 갑자기 소멸되지 않고 있음 을 알 수 있다. 이와 같은 결과는 계절변동의 불안정적 특성을 나타내는 것으로 따라서 계절 차분이 요구된다. 또한 <그림 6>의 편 자기상관 함수 PACF도 역시 "+"부호부터 시작하되 부호를 바꾸어 가며 수축적인 진동을 하며 계절주기에 해당되는 시차를 따라 스파이크를 기록하면서 소멸되지 않고 아주 느리게 감소하고 있음을 보여주고 있다. 아래<그림 7>은 1차 차분한 후의 자기상관 함수 ACF의 도표이다. lag 1부터 지수적으로 급격히 감소하고 있으며 또한 lag12 이후부터도 빠르게 0으로 감소하는 패턴을 보이고 있다. 특히 lag12에서는 음의 값을 갖는 돌출 형태를 띠고 있으며, <그림 7>의 추정된 ACF의 모형에서 AR(1)과정의 지수적으로 감소하는 자기상관 함수의 형태는 ACF 값들 중에서 lag1 이후의 형태와 lag12이후의 형태를 보면 짐작할 수 있으며 계절 AR(1)과정의 ACF 패턴은 lag12, 24에 해당하는 ACF 값을 보면 알 수 있다. 한편,<그림 8>의 편 자기상관함수 PACF 도표에서는 그 이외의 시차에 대해서는 별로 유의하지 않은 것으로 보여진다.

Name of Variable = wt

Period(s) of Differencing	1,12
Mean of Working Series	-0.00012
Standard Deviation	0.204116
Number of Observations	155
Observation(s) eliminated by differencing	13

Autocorrelations

Lag	Covariance	Correlation
0	0.041663	1.00000
1	-0.012929	-.31031
2	-0.0044953	-.10790
3	0.0062861	0.15088
4	-0.0093162	-.22365
5	0.00054131	0.01299
6	0.0034674	0.08322
7	-0.0037339	-.08962
8	0.0043902	0.10537
9	-0.0016753	-.04021
10	-0.0046762	-.11223
11	0.011500	0.27617
12	-0.016795	-.40311
13	0.0057580	0.13820
14	0.0050147	0.12036
15	-0.0049138	-.11794
16	0.0025885	0.06213
17	0.0016562	0.03975
18	-0.0068329	-.16400
19	0.00007812	0.00188
20	0.0020033	0.04808
21	-0.0012476	-.02995
22	0.0063080	0.15140
23	-0.0031175	-.07483
24	0.00050543	0.01213
25	-0.00033500	-.00804
26	-0.0023619	-.05673
27	0.0018030	0.04328
28	0.0017855	0.04286
29	-0.0061594	-.14784
30	0.0054659	0.13119
31	0.0010059	0.02414
32	-0.0060936	-.14626
33	0.0035472	0.08514
34	-0.0035564	-.08536
35	0.0012471	0.02993
36	0.0041204	0.09890

Partial Autocorrelations

Lag	Correlation
1	-0.31031
2	-0.22595
3	0.04631
4	-0.20560
5	-0.11804
6	-0.02898
7	-0.06447
8	0.03983
9	-0.04059
10	-0.11738
11	0.20501
12	-0.33172
13	0.00125
14	-0.03168
15	0.05147
16	-0.07184
17	-0.00220
18	-0.11652
19	-0.15393
20	-0.03366
21	-0.02796
22	0.01655
23	0.09611
24	-0.12288
25	-0.00330
26	-0.03747
27	0.03951
28	-0.01121
29	-0.13582
30	0.04997
31	-0.02018
32	-0.04967
33	-0.05469
34	-0.09155
35	0.05661
36	-0.01670

<그림 7> 시계열의 1차 차분후의 ACF <그림 8>시계열의 1차 차분후의 PACF

3.2.4 모형추정

<표 2> 최우추정 결과표

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
AR 1,1	-0.22654	0.07991	-2.84	0.0046	1
AR 2,1	-0.48758	0.07210	-6.76	<.0001	12

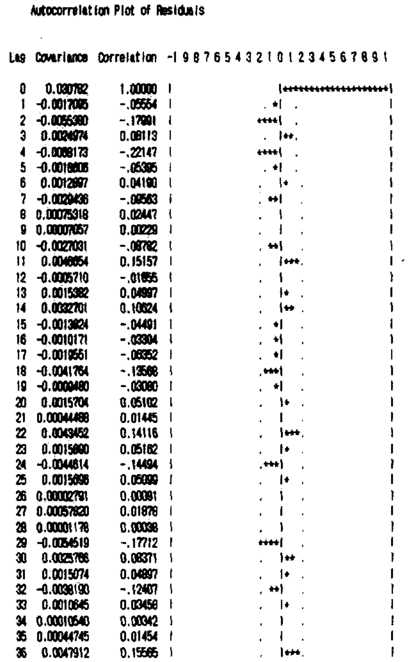
Variance Estimate	0.030782
Std Error Estimate	0.175447
AIC	-94.3617
SBC	-88.2748
Number of Residuals	155

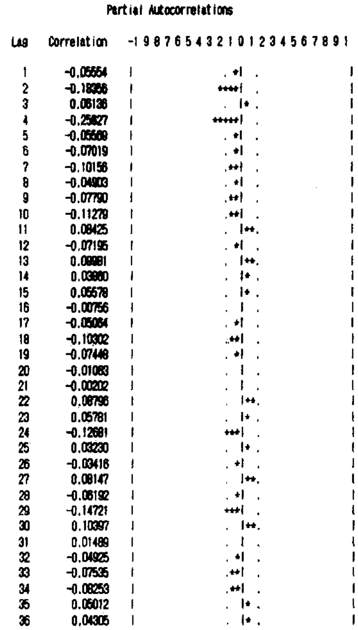
<표 2>의 최우 추정결과표와 같이 추정 계수들의 유의성 검정을 위해 산출한 t값들의 모수를 추정한 결과 유의수준 5% 에서 모두 유의한 값을 가지고 있음을 알 수 있다.

ARIMA 모형의 적합성을 알려주는 통계량이 가장 작은 값을 바탕으로 추정한 ARIMA모형은 $(1.0.0)\times(1.1.0)_{12}$의 모형을 따름을 알 수 있다. 최우법은 우도함수(likelihood function)에 자료가 포함하고 있는 모수(parameter)에 관한 모든 유용한 정보를 반영하기 때문에 바람직한 추정량의 통계적 속성을 만족시킨다. 그러나 기본속성상 비선형의 특성을 갖고 있는 ARIMA 모형의 정확한 최우 추정치를 찾는 것은 어려운 작업이 될 뿐만 아니라 상대적으로 많은 시간이 요구된다. 따라서 Box-Jenkins는 최소 자승법을 적용할 것을 권고하였다. 만일 모형의 확률충격(ε_t)이 정규분포를 따른다면 AR모형의 경우 최소 자승추정량은 최우추정량과 같거나 비슷하기 때문이다[9].

3.2.5 모형진단

추정된 ARIMA모형$(1.0.0)\times(1.1.0)_{12}$이 관측된 시계열에 잘 적합 된다면 이 모형은 시계열이 가지고 있는 모든 형태의 패턴을 다 설명하게 되므로 잔차에는 아무런 패턴이 포함되지 않는 순수한 오차의 특성인 백색잡음 과정의 성질을 갖게 된다. SAS를 통해얻어진 잔차의 ACF 와 PACF의 분석도표를 살펴보면 그 형태가 백색잡음 과정과 거의 동일한 결과를 얻어 모형 적합이 잘 되었음을 의미한다. 선택된 시계열 모형을 이용하여 미래 k시점의 시계열 값인 Z_{n+k} 을 예측하기 위해 SAS의 FORECAST를 이용하여 다음 시점의 값을 예측하였다. 이 값들은 X11-ARIMA 모형에서 결측 값들을 대신해 주어 X11-ARIMA모형의 정확성을 더욱 높여준다.

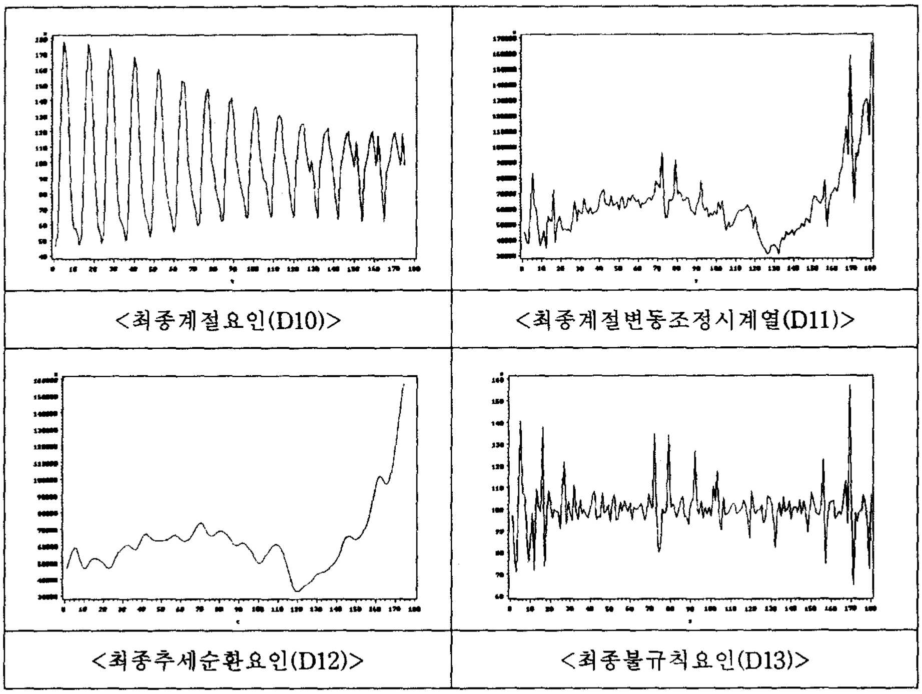




<그림 9> 추정된 모형에 의한 잔차분:ACF <그림 10> 추정된 모형에 의한 잔차분석:PACF

3.2.6 계절 조정 및 성분 분해

분해법은 시계열을 구성하는 각 성분들을 따로 구분한 후 이를 이용하여 미래를 예측하는 데 목적이 있다. 즉 추세성분은 어느 정도 일정한 비율로 증가 또는 감소하므로 단기적인 기간동안에는 이와 유사한 추세가 변함없이 계속되리라는 예상을 할 수 있다. 계절성분도 조금씩 변화하더라도 역시 단기간 내에서는 큰 변화가 없으리라고 생각할 수 있으므로 이러한 성분들을 따로 분해해서 생각한다면 좀더 예측이 정확하지 않을까 하는 것이 분해법의 목적이다. 또한 계절적인 성분을 시계열에서 제거한 후 장기적인 추세만을 분석하는 것이 중요한 경제관련 자료에서 계절 조정을 하기 위해서 분해법을 사용한다. 이렇게 분해법에 의해 추정된 계절성분을 원래의 시계열에서 제거한 시계열을 계절변동 조정 시계열이라고 한다. ARIMA모형을 통해서 구한 예측값을 바탕으로 X11-PROC를 사용하여 최종적으로 계절요인(D10), 계절변동조정 시계열(D11), 추세순환 요인(D12), 그리고 불규칙요인(D13)을 구한 후 그래프를 나타내면 다음 <그림 11>과 같다.



<최종계절요인(D10)>

<최종계절변동조정시계열(D11)>

<최종추세순환요인(D12)>

<최종불규칙요인(D13)>

<그림 11>최종계절변동조정 시계열의 요인별 분해Graph

4. 수요 예측

4.1. 수요예측방법

SAS 패키지에서 제공하는 X11 프로시져에는 ARIMA라는 옵션을 사용하여 모형을 자동 선택할 수 있다. 기본적으로 내장된 모형을 통해서 원시계열 자료에 가장 적합한 모형의 결과를 추출하는 것이다. 기본적으로 내장된 모형은 <표 3>과 같다.

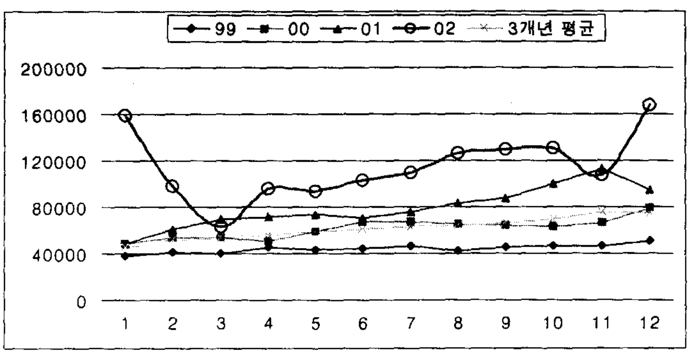
<표 3> 선정된 5가지 ARIMA모형

Model	Specification	Multiplicative	Additive
1	(0,1,1)(0,1,1)s	log transform	no transform
2	(0,1,2,)(0,1,1)s	log transform	no transform
3	(2,1,0)(0,1,1)s	log transform	no transform
4	(0,2,2)(0,1,1)s	log transform	no transform
5	(2,1,2)(0,1,1)s	no transform	no transform

<표 3>에 선정된 5가지의 모형은 캐나다에서 사용한 X11-ARIMA/88 모형과 동일하다. 하지만 자동모형을 사용하여 적합했을시에 유의수준에 들어오는 모형이 존재하지 않았다. 따라서 본 논문에서는 수동으로 ARIMA-PROC와 X11-PROC를 순차적으로 사용하여 SAS 패키지 상에서 X11-ARIMA 모형을 구현하였다. 먼저 X11-ARIMA 모형을 구현하기 위해서는 가지고 있는 시계열 자료에 가장 적합한 ARIMA 모형을 선정을 해야 한다. 모형 선정의 기준은 AIC 또는 SBC 값이 가장 작은 값을 가질수록 보다 유의한 모형이라고 볼 수 있다. 선정된 ARIMA모형을 통하여 데이터의 다음 연도 값을 추정하고 그 결과값과 함께 원 시계열을 X11-PROC를 이용하여 적합하면 다음 년도의 계절 조정된 시계열을 얻을 수 있다. 한편, X11-ARIMA 모형을 통해서 장기간의 예측값을 구하기 위해서는 매년 실제로 발생된 데이터를 업데이트 하여 모형에 추가하는 방식으로 반복해서 새로운 예측값을 추정할 수 있다. 기업은 매년 전년도 자료를 바탕으로 다음해 판매량을 예측하고 예산을 책정하게 된다. 정확한 예측을 위해서는 계절성으로 인한 불확실성을 제거하고 추세 성분만을 바탕으로 판매 신장률을 가늠해 볼 수 있다. 1988년 1월부터 2001년 12월까지의 14년(168개월)간의 자료를 바탕으로 2002년도 냉장고 수요를 예측하기 위해서는 자료의 계절적인 특성을 충분히 고려 해야한다. X11-ARIMA를 이용해서 구한 2002년도 수요의 예측값과 최근 3개년 판매실적을 평균한 실적치를 분석한 결과는 다음과 같다.

4.2. 최근 3년평균 실적치와 수요예측

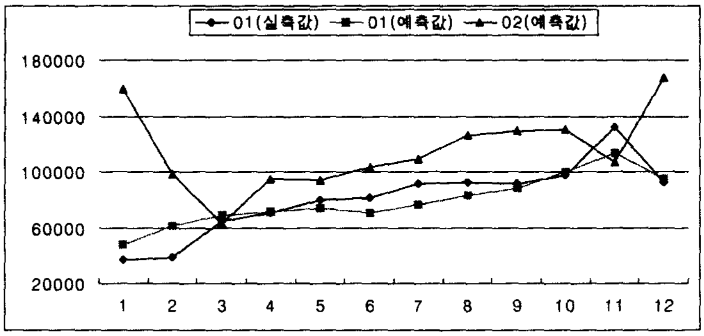
<그림 12>의 X11-ARIMA 모형에 의해서 수정된 최근 3개년 자료를 분석해 보면 월마다 불규칙하게 존재한 계절적 성분이 제거되고 99년부터 꾸준하게 판매량이 증가하고 있는 현상을 보여주고 있다. 양 끝 값은 이동평균과정에서 발생한 오차이다. 3월부터 11월까지 전년대비 판매량은 확실하게 증가하고 있음을 알 수 있다. 성수기인 여름보다 가을에 판매량이 증가한 이유는 기업의 판촉행사로 인한 마케팅 효과가 나타난 것으로 보인다. 이를 바탕으로 2002년도 냉장고 수요량을 예측할 수 있다.



<그림 12> 최근 3년평균실적치와수요예측

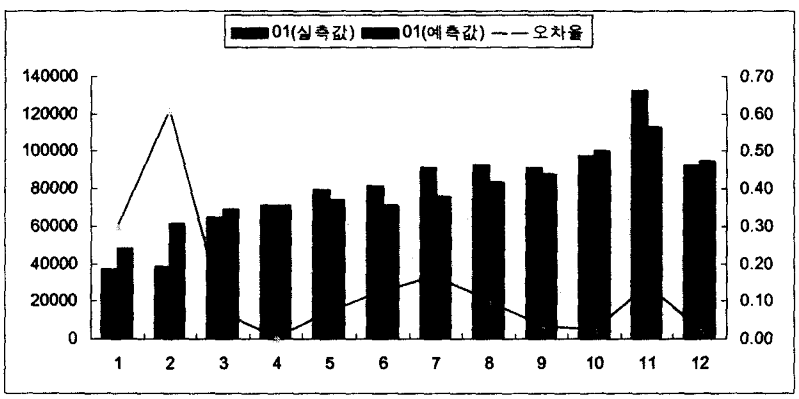
4.3. 2001년실적기준 수요예측

X11-ARIMA 모형으로 추정된 모델의 예측값과 실제 판매실적과의 차이 정도를 알아보기 위해서 2001년도 자료를 바탕으로 예측값과 실적치를 분석해본 결과 아래<그림 13>과 같다.



<그림 13> 수요예측

위<그림 13>그래프를 통해서 2001년도 실제 판매량과 모형에 의해서 수정된 예측값이 크게 차이가 나고 있지 않음을 알 수 있다. 2월과 11월달의 강한 계절성이 어느 정도 평활화 되었음을 볼 수 있다. 한편 추정된 2002년도 판매량은 전년도 보다 전반적으로 판매량이 증가할 것으로 추정하고 있다. 참고로 양끝값은 이동평균 과정에서의 오차값으로 인한 이상값이다.



<그림 14> 수요오차

한편, 예측값과 실측값의 오차율을 계산한 결과 <그림 14>와 같다. 2월의 높은 오차율은 계절적 요인, 판촉전략, IMF 영향등 98년이후 3년 평균값이 낮게 나타나서 14년간의 판측값을 이동평균하여 계절조정된 모형에 의한 예측값과 차이가 발생, 좋은 결과를 주지 못한 결과로 보여지며, 이런 특정 변수를 제외하고는 비교적 낮은 오차율 나타내고 있다.

5. 결론 및 추후연구과제

본 연구에서는 경제 시계열에 내재된 각 성분들을 따로 구분한 후 계절 조정을 하기 위해서 분해법 의해 추정된 계절성분을 원래의 시계열에서 제거하여 계절변동 조정하였으며. ARIMA모형을 이용하여 구한 예측값을 바탕으로 X11-PROC를 사용하여 계절조정 및 성분을 분해하여 최종적으로 계절요인, 추세순환요인, 불규칙인요인, 그리고 계절변동조정 시계열의 값을 구하는 과정과 방법을 제시하였다. 본 연구에서 계절변동조정에 적용된 X11-ARIMA 모형에서도 이동평균 연산을 반복하여 수행하는 과정에서 양 끝값이 이상 값을 가지는 경우가 많음을 발견하였다. 보다 정확한 모형추정을 위해서는 결측값 보정, 이상값 제거 등의 작업이 선행되어야 함을 뜻한다. 계절적 성분을 가진 시계열 자료에서 X11-ARIMA 모형을 이용하여 일반적인 경제, 경영 시계열에서 정확한 추세, 순환 현상을 예측할수 있음을 발견하였다. AR(Autoregressive)모형에 의하여 이 요인들을 제거하여 추세순환 요인, 계절요인 및 불규칙 요인을 조정한 방법에서는 주기 12인 함수에 의한 계절요인의 설명력이 X11-ARIMA 방법 보다 향상되었다. 경제와 경영 현상 속에서 발생하는 많은 정보들은 시계열 자료들이다. 특히 일년을 주기로 계절성을 가진 자료들이 많이 있다. 하나의 정보는 많은 요인들에 의해서 끊임없이 변화고 있다. 본 연구에 적용한 X11-ARIMA모형은 수많은 불확실성 중에서 계절적인 요인을 제거해 보다 정확한 시계열 자료를 얻을 수 있다는 점에서 그 효용성이 있다 하겠다. 기업은 매년 전년도 자료를 바탕으로 다음해 판매량을 예측하고 예산을 책정하게된다. 정확한 예측을 위해서는 계절성으로 인한 불확실성을 제거하고 추세 성분만을 바탕으로 판매량, 신장률. 생산량, 재고량 등을 가늠 볼 수 있다. 본 연구에서는 계절조정 결과를 향상시키기 위하여 X11-ARIMA 방법에 기초한 적합한 모형에 의해 수요예측을하여 자기상관과 계절성을 고려한 예측방법을 제시하였다. 수요예측은 과거의 판매량만을 기초로 한 단일변량 시계열분석(Univariate Time Serires Analysis)으로 행하여 졌다. 본 연구의 결과를 바탕으로 판매량 이외에 신상품, 가격, 판촉전략 등을 고려한 추가적인 연구가 진행 될 수 있을 것으로 예상된다.

참고문헌

[1] 김덕(1992), 「계량경제학」, 인간사랑.

[2] 김동호(1989), 한국의 통화량 시계열의 계절조정, 연세대 석사학위 논문.

[3] 김태완(1985), 시계열분석에 따른(O · R) 재고통제 모형에 관한 연구, 연세대 대학원 석사 학위논문.

[4] 박영진(1994), 계절성 결정적 추세를 갖는 시계열 모형에서의 단위근 검정법, 서울대

대학원 박사학위논문.
[5] 박유성, 허명회(2000), 시계열분석자료, 자유아카데미.
[6] 박연서(2000), 신상품 수요예측모형 개발 및 통신시장에의 응용, 한국과학기술원 박사학위논문.
[7] 신호중(1996), 한국경제의 경기순환예측을 위한 통계적 방법에 관한 연구, 동국대대학원 박사학위논문.
[8] 염준근(2000), 「선형회귀분석」,자유아카데미.
[9] 이종원(1994), 「계량경제학」, 박영사.
[10] 장경수(1997), 수요예측이 재고관리정책에 미치는 영향에 관한연구, 연세대대학원 박사학위 논문.
[11] 전태준(1991), 계절적 시계열 모형화를 위한 VSACF의 확장, 한국경영과학회지, vol.16. No.1.
[12] 주영진(1995), 장ㆍ단기 요인의 분리를 위한 시계열 모형과 그 응용: 경기순환변동 예측 및 신상품 수요예측, 한국과학 기술원 박사학위논문.
[13] 한국은행(1989), X-11 ARIMA방법에 의한 분기 GNP 계절변동조정.
[14] Akaike, H.(1974), A New look at statistical model identification, *IEEE Trans, on Automatic Control,* vol.19, pp716-723.
[15] Box(1976), G. E. P. and Jenkins, G. M., *Time Series Forecasting and Control,* Holden Day, NewYork,.
[16] Bureau of the Census.X-11 Information for the User, *U. S. Department of Commerce* Government Printing Office, Washington D. C, (1969).
[17] Chou, K., Higginson, J. and Hout, G.(1985), Perfomance of ARIMA Models in Time Series, *Survey Methodology,* vol.11, pp51-64, .
[18] Cleveland, W. P. and Dempster, A. P., Advances in model-based seasonal adjustment, *ADA Proceedings in Business and Economics,* (1980).pp30-37.
[19] Dagum, E. B.(1988) *The X11ARIMA/88 Seasonal Adjustment Mehtod:* Foundations *and Users Manual,* Canada.
[20] Hiller, S. C and Ciao, and G. G.,(1982), An ARIMA model-based approach to seasonal adjustment, *JANA,* 77(377), pp63-70,
[21] Lothian, J and Morry(1978), M. *A Set of Quality Control Statistics for The X11-ARIMA Seasonal Adjustment Program* Canada.
[22] Markridakis, S and Wheelwright, S. (1978), *Forecasting : Methods and Application,* John Wiley and Sons, Inc. New York,.
[23] Pandit, S. M., *"Data dependent system approach to trend and seasonality"* Time Series Analysis : Theory and Practice 1, Anderson, O. D.(ed), pp515-526, 1982.
[24] Pandit, S. M. and We, S, M.(1983), *Time Series and Systems Analysis,* John Wiley and Sons, Inc. New York.

[25] Walter Anders Iowa State Univ, *Applied Economic Time Series*, John Wiley & Sons, Inc, New York, 1995.
[26] Winters, P, R,(1960), Forecasting sales by exponentially weighted moving average, *Management Science,* vol.6, pp324-42.