

A study on parsimonious periodic autoregressive model

Jiho Lee^a · Byeongchan Seong^{a,1}

^aDepartment of Applied Statistics, Chung-Ang University

(Received December 15, 2015; Revised January 2, 2016; Accepted January 2, 2016)

Abstract

This paper proposes a parsimonious periodic autoregressive (PAR) model. The proposed model performance is evaluated through an analysis of Korean unemployment rate series that is compared with existing models. We exploit some common features among each seasonality and confirm it by LR test for the parsimonious PAR model in order to impose a parsimonious structure on the PAR model. We observe that the PAR model tends to be superior to existing seasonal time series models in mid- and long-term forecasts. The proposed parsimonious model significantly improves forecasting performance.

Keywords: seasonal time series model, parsimony of principle, seasonality, ARIMA model, Holt-Winters model, unemployment rate

1. 서론

대부분의 시계열 자료, 예를 들어 국내총생산(GDP), 실업률, 주택매매가격, 소비지출 등의 경제지표들을 보면 계절에 따라 일정한 패턴이 반복되는 경향을 가지고 있다. 전통적인 분석에서는 사전 계절조정을 통해 계절에 대한 변동을 제거하여 모형화하는 것이 일반적이지만, 이러한 계절조정이 정보손실을 유발할 수 있다는 연구 결과들이 발표되어 왔다; 예를 들면, Franses와 McAleer (1998), Matas-Mir와 Osborn (2004)를 참고하여라. 이에 따라 계절적 변동을 동시에 모형에 고려하기 위한 계절형 모형에 대한 이론적, 실증적 연구가 활발히 진행되고 있다.

계절을 모형화하기 위한 많은 연구가 있었고 그 중 대표적인 모형이 Box와 Jenkins (1976)의 SARIMA (seasonal autoregressive integrated moving average) 모형이다. 또다른 계절 시계열 모형으로 시계열 분해법을 통해 수준(level), 추세(trend), 계절(seasonality)로 모형화하는 Holt-Winters 모형이 있으며 이외에도 계절 시계열을 모형화하기 위한 많은 모형들이 연구되어 왔다.

본 논문에서 소개하는 PAR(periodic autoregressive) 모형은 계절에 따른 변동을 계절별로 다르게 반영할 수 있는 모형으로서, 계절에 의존하는 모수를 고려하여 모형화한다. 주기적 시계열 모형(periodic times series model)의 필요성은 Pagano (1978)에 의해 제기되었으며 Osborn과 Smith (1989)에서 경제 문헌에 소개되었다. 그 이후 연구에서 McLeod (1993)는 낮은 차수의 PAR 모형만으로도 주기적 시계열을 설명하기 충분하다는 것을 밝혔으며, Osborn (1991)은 주기성을 무시하면 계절조정 후에도 주기적 변동을 가지게 되어 예측치에 편이가 생김을 보였다. 최근에는 최적의 PAR 차수를 찾기 위해

¹Corresponding author: Department of Applied Statistics, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Korea. E-mail: bcseong@cau.ac.kr

유전자 알고리즘을 이용하는 방법 (Ursu와 Turkman, 2012)이 연구되고 있으며 국내 연구로 Lee 등 (2012)은 PAR 모형의 확장된 형태인 PARMA 모형을 이용하여 Arosa 지방의 성층권 오존 시계열에 대한 분석에서 SARIMA 모형보다 주기적 모형이 더 우수함을 보였다. 그러나 PAR 모형은 일반적으로 계절에 따라 모수의 수가 빠르게 늘어나는 문제점을 가지고 있으며, 이에 대한 개선을 위한 연구는 Lund 등 (2006)의 연구 이외에는 부족한 실정이다.

본 논문에서는 모수의 수가 많은 PAR 모형에 적절한 제약식을 부여하고 이에 대한 LR 검정을 활용하여 모수 절약 모형에 대한 연구를 하였다. 또한, 실업률 계절 시계열 자료 분석을 통하여 모수 절약 PAR 모형과 SARIMA 및 Holt-Winters 모형간의 예측력을 비교하였다.

본 논문은 총 5장으로 구성되어 있으며 2장에서는 SARIMA와 Holt-Winters모형을 소개하고 3장에서는 PAR 모형의 형태, 추정방법 및 모수 절약 모형, 주기적 적분(periodic integration)을 설명한다. 4장에서는 국내 실업률 자료를 이용하여 모형의 적합 및 예측성능을 비교분석하며 5장에서 결과를 정리하였다.

2. 전통적 계절 시계열 모형

2.1. SARIMA 모형

SARIMA 모형은 Box와 Jenkins (1976)에 의해 제안된 모형으로 계절 변동을 갖는 시계열을 모형화 하기 위하여 가장 널리 사용되고 있다. 시계열이 정상성(stationarity)을 만족하지 않는 경우 비정상성을 제거하기 위해 차분을 이용하여 정상성을 만족시킨 후 ARMA 모형에 적합하게 된다. 시계열 y_t 가 $ARIMA(p, d, q) \times (P, D, Q)_s$ 를 따르면 아래와 같이 표현할 수 있다,

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t. \quad (2.1)$$

이때,

$$\begin{aligned} \phi_p(B) &= 1 - \phi_1 B - \dots - \phi_p B^p, & \Phi_P(B^s) &= 1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}, \\ \theta_q(B) &= 1 - \theta_1 B - \dots - \theta_q B^q, & \Theta_Q(B^s) &= 1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs} \end{aligned}$$

를 나타낸다. 단, p, d, q 는 각각 AR의 차수, 차분차수, MA의 차수를, P, D, Q 는 계절 AR의 차수, 계절 차분차수, 계절 MA 차수를 나타내며; B 는 후진연산자(back shift operator), s 는 계절주기, ε_t 는 오차항으로 평균이 0이고 분산이 일정한 백색잡음과정(white noise process)이다.

2.2. Holt-Winters 모형

Winters (1960)의 가법적 계절 지수평활법(additive seasonal exponential smoothing)은 Holt (1957)의 이중 모수 지수평활법을 계절형으로 확장시킨 모형으로 계절성이 추세에 영향을 받지 않는 경우 사용하는 모형이다. Holt-Winters 모형은 시계열의 움직임을 수준, 추세, 계절성분으로 분해하여 모형화하며 1시점 이후의 예측값인 \hat{y}_{t+1} 는 다음과 같이 표현된다,

$$\hat{y}_{t+1} = l_t + b_t + s_{t+1-p}.$$

이 때, l_t, b_t, s_t 는 각각 수준성분, 추세성분, 계절성분을 의미하며 아래와 같이 정의된다;

$$l_t = \alpha(y_t - s_{t-p}) + (1 - \alpha)(l_{t-1} + b_{t-1}),$$

$$\begin{aligned} b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}, \\ s_t &= \gamma(y_t - l_t) + (1 - \gamma)s_{t-p}, \end{aligned}$$

여기서 y_t 는 t 시점에서의 관측값, p 는 계절 주기를 나타내고, α, β, γ 는 각각 수준, 추세, 계절성분에 대한 평활상수를 의미하며 $0 \leq \alpha, \beta, \gamma \leq 1$ 를 만족한다.

3. PAR 모형

본 장에서는 주기적 자기회귀(PAR) 모형을 설명하며 주기적 적분에 대한 개념을 살펴본다. 더 자세한 설명은 Franses와 Paap (2004)를 참고하여야. PAR 모형은 계절 시계열 모형으로서 주로 분기별(quarterly) 자료나 월별(monthly) 자료에 대한 모형화에 이용된다. 본 논문에서는 간단하게 분기별 자료에 대한 PAR 모형을 설명한다.

3.1. 기본 모형

본 모형은 계절성을 가지는 자료에 대하여 모형의 계수(coefficient) ϕ 들이 주기적으로 변하는 구조를 가지는 모형으로, 다음과 같은 p 차의 PAR 모형, $\text{PAR}(p)$, 형태로 나타낼 수 있다;

$$y_t = \phi_{1,s}y_{t-1} + \cdots + \phi_{p,s}y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, 1). \quad (3.1)$$

단, $t = 1, 2, \dots, n = 4N$, $s = 1, 2, 3, 4$ 이며, y_t 는 분기별로 N 년간 관측된 자료에서 각 계절 평균(μ_s)을 제거한 값으로 계절(s)에 따라 각각 N 개의 값이 존재한다. 따라서, 각 계절에 따라 다른 PAR 모형의 계수값을 가진다. 이로 인하여 연단위(annual observation)로 모형을 재표현할 수 있으며 PAR 모형은 각 계절에 대한 4개의 다른 $\text{AR}(p)$ 모형으로 구성되므로 4차원의 벡터 $\text{AR}(p)$ 모형 형태로 표현 가능하다. 자세한 내용은 Osborn (1991)과 Lütkepohl (2005)을 참고하여야,

$$\Phi_0 Y_T = \Phi_1 Y_{T-1} + \cdots + \Phi_P Y_{T-P} + \varepsilon_T. \quad (3.2)$$

단, $T = 1, 2, \dots, N$, $P = 1 + [(p-1)/4]$ 이며; $[x]$ 는 x 의 정수부분을 나타낸다.

$$\begin{aligned} Y_T &= (Y_{1,T}, Y_{2,T}, Y_{3,T}, Y_{4,T})', \\ \varepsilon_T &= (\varepsilon_{1,T}, \varepsilon_{2,T}, \varepsilon_{3,T}, \varepsilon_{4,T})' \end{aligned}$$

는 4차원의 벡터 형태이며 $Y_{s,T}$ 는 계절 s 에서 T 시점의 관측값, $\varepsilon_{s,T}$ 는 계절 s 에 대한 T 시점의 오차를 나타낸다. $\Phi_0, \Phi_1, \dots, \Phi_P$ 는 (4×4) 차원의 모수 행렬이며 아래와 같이 표현된다;

$$\Phi_k(i, j) = \begin{cases} 1, & i = j, \\ 0, & j > i, \\ -\phi_{i-j,i}, & j < i, \end{cases} \quad \Phi_k(i, j) = \phi_{i+4k-j,i}.$$

단, $i, j = 1, 2, 3, 4$ 이고 $k = 1, 2, \dots, P$ 이다. 예를 들어, $\text{PAR}(2)$ 모형의 경우에 다음과 같이 단변량 형태인 식 (3.3)과 벡터 형태인 식 (3.4)로 표현 가능하다.

$$y_t = \phi_{1,s}y_{t-1} + \phi_{2,s}y_{t-2} + \varepsilon_t, \quad (3.3)$$

$$\Phi_0 Y_T = \Phi_1 Y_{T-1} + \varepsilon_T. \quad (3.4)$$

이때, Φ_0 와 Φ_1 은 아래와 같이 표현된다;

$$\Phi_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\phi_{1,2} & 1 & 0 & 0 \\ -\phi_{2,3} & -\phi_{1,3} & 1 & 0 \\ 0 & -\phi_{2,4} & -\phi_{1,4} & 1 \end{pmatrix}, \quad \Phi_1 = \begin{pmatrix} 0 & 0 & \phi_{2,1} & \phi_{1,1} \\ 0 & 0 & 0 & \phi_{2,2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

단변량 형태인 식 (3.1)을 정리하면 다음과 같은 형태로 표현이 가능하며 이러한 인수 분해된 형태(roots of the factorized AR)를 주기적 차분형태(periodically differenced form)라고 한다.

$$y_t - \alpha_s y_{t-1} = \beta_{1,s}(y_{t-1} - \alpha_{s-1}y_{t-2}) + \cdots + \beta_{(p-1),s}(y_{t-(p-1)} - \alpha_{s-(p-1)}y_{t-p}) + \varepsilon_t. \quad (3.5)$$

3.2. 주기적 적분

식 (3.2)를 이용한 다음의 특성 방정식 해가 단위근 밖에 존재하는 경우 주기적 정상(periodic stationary)이라고 정의할 수 있다,

$$\left| \Phi_0 - \Phi_1 L - \cdots - \Phi_p L^p \right| = 0.$$

확률적 추세가 시계열 자료에 존재하는 경우 특성 방정식의 근이 단위근 위에 존재하며 이를 주기적 적분 계열(periodically integrated series)이라 한다. 예를 들어 y_t 가 분기별 자료이며 주기적 적분 계열 일 때 확률적 추세를 제거하기 위해 차분 필터인 $(1 - \alpha_s B)$ 가 필요하다. 이때, 식 (3.5)의 모수 α_s 는 $\alpha_1 \alpha_2 \alpha_3 \alpha_4 = 1$ 과 모든 $s = 1, 2, 3, 4$ 에 대해 $\alpha_s \neq \alpha$ 를 만족하며 이를 1차 주기적 적분(periodically integrated of order 1; PI(1))이라고 한다. 자세한 내용은 Franses와 Paap (2004)를 참고하여라.

3.3. 모수 추정 방법

PAR(p) 모형은 계절 s 인 경우 1값을 가지고 그 외에는 0값을 가지는 계절형 가변수 $D_{s,t}$ 를 이용하여 아래와 같은 회귀 모형식으로 표현 가능하다,

$$y_t = \left(\sum_{s=1}^4 \phi_{1,s} D_{s,t} \right) y_{t-1} + \cdots + \left(\sum_{s=1}^4 \phi_{p,s} D_{s,t} \right) y_{t-p} + \varepsilon_t. \quad (3.6)$$

오차항 ε_t 가 정규분포로 가정될 때, 모수 $\phi_{i,s}$ ($i = 1, 2, \dots, p; s = 1, 2, 3, 4$)의 최대우도추정량은 식 (3.6)의 최소제곱추정법과 동일한 방법으로 쉽게 계산할 수 있다. 이외의 추정방법과 점근적 성질에 관해서는 Troutman (1979)를 참고하여라. 추가적으로, PAR 모형을 추정하기 위한 표본의 길이는 최소 $N = n/4$ 이 되어야 한다.

3.4. 모수 절약 PAR 모형

본 연구에서는 PAR 모형에 모수 절약 구조를 반영하기 위하여 식 (3.1)을 아래와 같은 형태로 고려할 수 있음을 제안한다,

$$y_t = \phi_{1,v} y_{t-1} + \cdots + \phi_{p,v} y_{t-p} + \varepsilon_t. \quad (3.7)$$

단, $t = 1, 2, \dots, n = 4N$ 이며 주기 v 는 원시계열의 주기 s 보다 작은 값을 가진다. 이러한 주기 v ($< s$)를 찾는 방법으로는, (i) 주기 s 에서 추정된 PAR 모형의 모수 추정 결과를 통하여 제약식의 형태를 찾

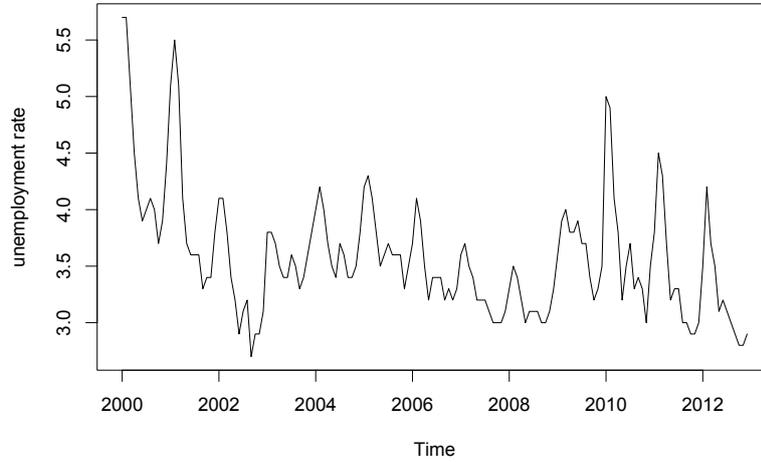


Figure 4.1. Times series plot of Korean unemployment rate.

거나; (ii) 계절별 시계열 그림을 통하여 계절들 간의 공통된 패턴을 이용할 수도 있으며; (iii) 흔히 믿어지는 계절들 간의 공통된 가설의 형태를 고려할 수 있을 것이다.

제약이 없는 PAR 모형(주기 s 를 가지는 모형)과 모수 제약 PAR 모형(주기 v 를 가지는 모형)간의 통계적 차이 유무는 다음과 같은 LR 검정통계량으로 검정할 수 있다,

$$LRT = -2[\ln(L_{RM}) - \ln(L_{FM})].$$

이때, L_{RM} 와 L_{FM} 는 각각 식 (3.7)의 모수가 제약된 모형과 식 (3.1)의 일반적인 PAR 모형의 로그 우도(likelihood)값을 나타내며, LRT는 모수 제약의 귀무가설 하에서 근사적으로 카이제곱분포를 따르며 자유도는 제약된 모수의 개수이다.

4. 실증분석

4.1. 자료

실증분석에 사용된 자료는 실업률 지수이며, 이 지수는 만 15세 이상의 인구 중에서 노동 의지와 능력은 있으나 일자리가 없어 실업 상태에 놓인 사람들의 비율로, $[\text{실업자} \div \text{경제활동인구}] \times 100(\%)$ 으로 계산된다. 본 연구에서는 2000년 1월부터 2015년 7월까지 총 187개의 월별 자료를 이용하였다. 2011년 12월까지의 모형적합을 위하여 사용하였고 이후의 자료는 예측의 성능을 평가하기 위하여 사용하였다. 이 자료는 통계청을 통해서 얻을 수 있다.

Figure 4.1은 실업률의 시계열 그림이다. 전반적으로 3%에서 5%사이에서 증가와 감소를 반복하는 주기적 패턴을 보이고 있다. 실업률의 주기를 세부적으로 확인하기 위해 월별 실업률을 1년 단위로 겹쳐서 그리면 Figure 4.2와 같다. 전반적으로 연초인 1월에 높았다가 시간이 지남에 따라 급격히 감소하였다가 다시 6월과 7월쯤에 증가하였다가 감소하고 연말에 다시 증가하는 패턴을 보이고 있다. 따라서 실업률 자료의 분석에서 계절성뿐만 아니라 계절(월) 사이의 공통된 특징이 고려되는 것이 적절하다고 보여진다. 본 논문에서는 2장과 3장에서 소개된 SARIMA 모형, 계절지수평활법(Holt-Winters), PAR 모형을 통해 분석 및 비교하고자 한다.

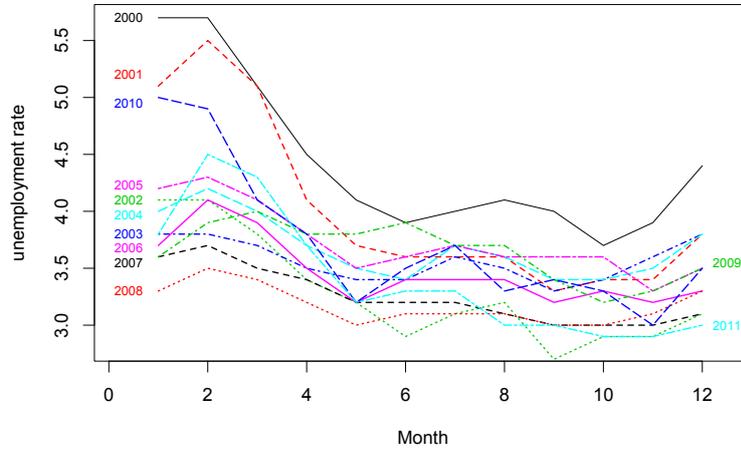


Figure 4.2. Month plot of Korean unemployment rate from 2000 to 2011.

Table 4.1. Estimation results for SARIMA and Holt-Winters models

SARIMA		Holt-Winters	
ϕ_1	0.6212	α	0.8898
θ_1	-0.7909	β	0.0001
Θ_1	-0.7122	γ	0.0001

SARIMA = seasonal autoregressive integrated moving average.

4.2. SARIMA 및 Holt-Winters 모형 적합

실업률 자료에 SARIMA 모형을 적합하기 위해 단위근 검정(augmented Dickey-Fuller test), ACF, PACF의 결과를 확인하였으며 최종적으로 AIC 기준으로 $ARIMA(1, 1, 1) \times (0, 1, 1)_{12}$ 모형을 적합하였다. Holt-Winters 모형은 시계열의 변동폭이 시간의 흐름에 관계없이 일정하므로 가법적 모형을 사용하였으며 모수를 추정하는데 있어 $MSE = n^{-1} \sum_{t=1}^n (y_t - \hat{y}_t)^2$ 값이 가장 작아지게 하는 α, β, γ 를 추정하였다. Table 4.1는 SARIMA와 Holt-Winters의 모수 추정 결과이다.

4.3. PAR 모형 적합

PAR 모형에 적합하기 위해서는 먼저 모형의 차수 p 를 결정해야 한다. 본 논문에서는 적절한 차수를 결정하기 위해 AIC, BIC를 기준으로 비교하였으며 보조적인 방법으로 F -test를 사용하였다. AIC와 BIC는 값이 작을수록 적절한 모형으로 판단할 수 있으며 F -test는 식 (3.6)에서 모든 주기에 대하여 p 차수까지의 모든 모수 $\phi_{p,s}$ 가 0이 아니면서, $p+1$ 번째 차수의 모수 $\phi_{p+1,s}$ 가 0이 되는지를 검정하는 것이다. 자세한 내용은 Franses와 Paap (2004)를 참고하여야. Table 4.2는 PAR 모형의 차수에 대한 AIC, BIC, F -test의 결과이다. 이 3가지 기준에 대한 결과를 종합하여 본 논문에서는 PAR(1) 모형을 최종모형으로 적합하였다.

주기적 정상성을 판단하기 위하여 Boswijk 등 (1997)이 고려한 LR 검정을 사용하였으며 검정 결과 주기적 정상성이 있다고 판단하였다. 이 때 사용된 검정통계량은 다음과 같다,

$$LR_{PI} = n[\ln(SSR_0) - \ln(SSR_a)]. \quad (4.1)$$

Table 4.2. Result of AIC, BIC and F -test for choosing the proper order p

Criterion	Periodic autoregressive order			
	1	2	3	4
AIC	-65.4004	-50.7108	-60.5912	-66.5649
BIC	-25.8359	25.2130	51.5350	81.6053
$F(\phi_{p+1,s} = 0)$	0.6194	2.0274	1.3561	5.4103
p -value	0.8227	0.0275	0.1992	0.0000

Table 4.3. Monthly estimation results for PAR model

Month	$\hat{\phi}_{1,s}$	Month	$\hat{\phi}_{1,s}$
1	1.0128 (0.1416)	7	0.8994 (0.1860)
2	0.8598 (0.0752)	8	0.9875 (0.1837)
3	0.7509 (0.0816)	9	0.9286 (0.1671)
4	0.6004 (0.1028)	10	0.7696 (0.1601)
5	0.8196 (0.1599)	11	0.9792 (0.1912)
6	0.7869 (0.1706)	12	1.2440 (0.1696)

Note: Values in parenthesis denote standard errors.

PAR = periodic autoregressive.

SSR_0 와 SSR_a 는 각각 $PAR(p)$ 모형에서 주기적 적분이 제약이 된 경우와 제약되지 않은 경우의 잔차 제곱합을 나타낸다.

Table 4.3는 $PAR(1)$ 으로 적합된 모형의 추정결과이다. 모수 추정값($\hat{\phi}_{1,s}$)을 보면 12월에 1.2440로 높은 값이 나타나고 있으며 7월, 8월, 9월에서도 대략 0.9정도로 다른 월에 비해 상대적으로 높은 값을 나타내고 있다. 이와 반대로 실업률이 급격하게 떨어지는 기간인 3월에서 4월 기간에는 상대적으로 낮은 모수 추정값인 0.6004와 0.7509로 나타나고 있다. 이는 Figure 4.2의 결과와 거의 유사한 패턴을 보이고 있으며 모든 모수들이 유의하다는 것을 확인할 수 있다. 이는 PAR 모형의 적절성과 계절들 간의 공통된 특징이 존재함을 보여주고 있다.

4.4. 모수 절약 PAR 모형 적합

본 절에서는 4.3절에서 제안된 PAR 모형에서 두 가지 모수 절약 PAR 모형을 고려하였다. 이때, 모형 I은 4.3절에서 추정된 $PAR(1)$ 로 제약이 없는 모형이며 모형 II은 분기별로 같은 주기를 가질 것을 가정한 모형으로 귀무가설

$$H_0 : \phi_{1,1} = \phi_{1,2} = \phi_{1,3} \ \& \ \phi_{1,4} = \phi_{1,5} = \phi_{1,6} \ \& \ \phi_{1,7} = \phi_{1,8} = \phi_{1,9} \ \& \ \phi_{1,10} = \phi_{1,11} = \phi_{1,12}$$

를 만족하는 모형이며, 모형 III은 Figure 4.2의 결과에 근거해 실업률이 증가하는 구간과 감소하는 구간별로 같은 주기를 가질 것을 가정한 모형으로

$$H_0 : \phi_{1,11} = \phi_{1,12} = \phi_{1,1} \ \& \ \phi_{1,2} = \phi_{1,3} = \phi_{1,4} \ \& \ \phi_{1,5} = \phi_{1,6} = \phi_{1,7} \ \& \ \phi_{1,8} = \phi_{1,9} = \phi_{1,10}$$

를 만족하는 모형으로 설정하였다.

Table 4.4는 세가지 PAR 모형(모형 I, 모형 II, 모형 III)에 따른 AIC, BIC 및 LR 검정의 결과이다. AIC, BIC 기준으로 일반적인 PAR 모형보다 모수 절약 PAR 모형이 더 적절한 모형으로 나타났으며, LR 검정 결과 역시 모수 절약 PAR 모형과 제약이 없는 PAR 모형간에는 통계적으로 유의한 차이가 나타나지 않았다.

Table 4.4. Result of AIC, BIC and LR test for three types of PAR models

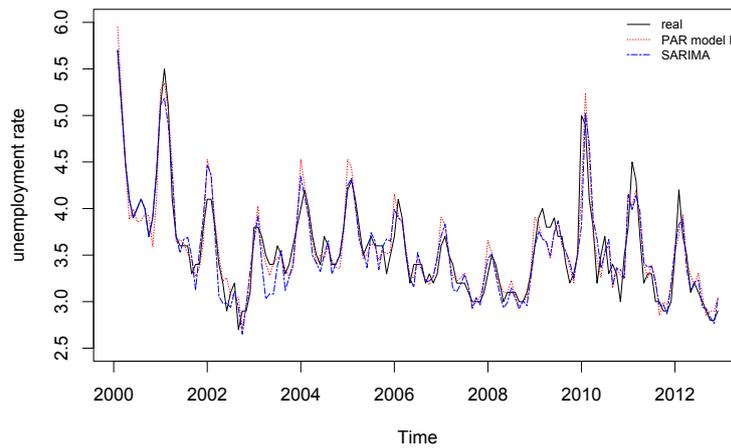
Criterion	PAR models		
	Model I	Model II	Model III
AIC	-65.4004	-72.1991	-74.2754
BIC	-25.8359	-56.9820	-59.0582
LR (<i>p</i> -value)	.	9.2013 (0.3256)	7.1251 (0.5232)

Note: The LR statistics test for significance of the corresponding model, compared with model 1.
 PAR = periodic autoregressive.

Table 4.5. Comparison of performances of the three models in the in-sample

In-sample	PAR models			SARIMA	Holt-Winters
	Model I	Model II	Model III		
RMSE	0.1912	0.1877	0.1887	0.2024	0.1947

PAR = periodic autoregressive, SARIMA = seasonal autoregressive integrated moving average,
 RMSE = root mean square error.

**Figure 4.3.** Comparison of performances of the three models in the in-sample.

4.5. 모형 비교 및 예측 성능 평가

본 절에서는 2000년 1월부터 2011년 12월까지 총 156개의 자료를 통해 PAR(1), 모수 절약 PAR(1), ARIMA(1, 1, 1) × (0, 1, 1)₁₂, Holt-Winters 모형을 적합하여 표본 내(in-sample)에서 평가함과 동시에, 2012년 1월부터 2015년 7월까지의 31개의 자료에 대하여 1개월, 6개월, 1년 이후의 예측을 통해 모형의 성능 평가를 실시하였다. 평가 기준으로는 실제값과 추정값의 차이를 나타내는 손실함수로 RMSE(root mean square error)를 이용하였으며 다음과 같다,

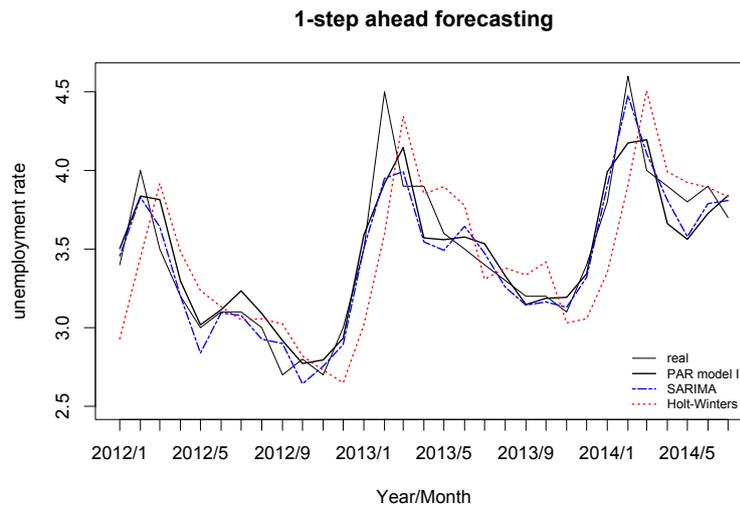
$$\text{RMSE} = \sqrt{m^{-1} \sum_{t=1}^m (y_t - \hat{y}_t)^2},$$

여기서 m 은 예측의 총 횟수를, \hat{y}_t 는 t 시점에서의 실업률의 추정값을 나타낸다. Table 4.5는 표본 내 자

Table 4.6. 1-, 6-, 12-step-ahead forecasting performances of the three models in the out-sample

Horizon	PAR models			SARIMA	Holt-Winters
	Model I	Model II	Model III		
1	0.2086	0.2148	0.1976	0.1575	0.3443
6	0.2874	0.2807	0.2896	0.2963	0.4182
12	0.2423	0.2451	0.2417	0.4690	0.5139

PAR = periodic autoregressive, SARIMA = seasonal autoregressive integrated moving average.

**Figure 4.4.** 1-step-ahead forecasting performances of the three models in the out-sample.

료에서의 5가지 모형에 대한 평가 결과이다. 표본 내에서는 전반적으로 PAR 모형의 RMSE값이 작게 나타나고 있으며, 모수 절약 모형일 때 성능이 더욱 개선되고 있다. Figure 4.3은 표본 내 자료에서의 실제 실업률과 추정값을 그린 것으로 편의상 가장 좋은 성능을 보이는 PAR 모형 II와 기존 계절 시계열 모형에서 가장 좋은 성능을 보인 SARIMA 모형을 나타낸 것이다. 전반적으로 실제 실업률과 비슷한 흐름을 보이고 있음을 확인할 수 있다.

Table 4.6은 1개월, 6개월, 1년 이후 표본 외 실업률의 예측에 대한 RMSE 결과이다. Figure 4.4~Figure 4.6은 각각 3개의 미래시점 예측에 대한 실제값과 예측값에 대한 그림으로 편의상 전반적으로 예측 성능이 가장 우수한 PAR 모형 III과 비교 대상으로 SARIMA, Holt-Winters 세가지 예측값만 나타내었다.

먼저 1개월 후 실업률 예측에 있어서는, SARIMA 모형의 예측력이 가장 좋았으며 PAR, Holt-Winters 모형 순으로 예측력이 좋게 나타났다. PAR 모형 중에서는 모수 절약 모형인 모형 III의 성능이 가장 좋게 나타났다. 6개월 후 예측 결과 PAR 모형일 때 좋은 예측력을 보이고 있으며, SARIMA, Holt-Winters 순으로 예측력이 좋게 나타났다. 이 경우에는 모수 절약 모형인 모형 II가 가장 좋은 예측 성능을 나타내었으나 전반적으로 큰 차이가 없는 것으로 보인다. 12개월 후 결과에서는, PAR 모형의 경우 6개월 이후 예측에 비해 예측력이 더 좋게 나타났으며 SARIMA와 Holt-Winters의 경우에는 예측력이 나빠졌다. 두 모형 모두 예측기간이 증가할수록 예측력이 급격하게 떨어지고 있음을 확인할 수 있다. 반면에 PAR 모형은 예측 기간에 상관없이 안정적인 예측력을 보이고 있으며 이는 Figure 4.6을 통해서도 확인할 수 있다.

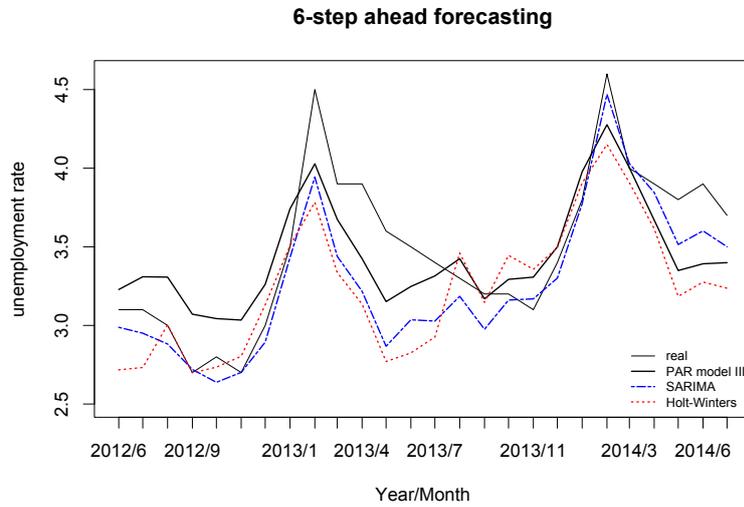


Figure 4.5. 6-step-ahead forecasting performances of the three models in the out-sample.

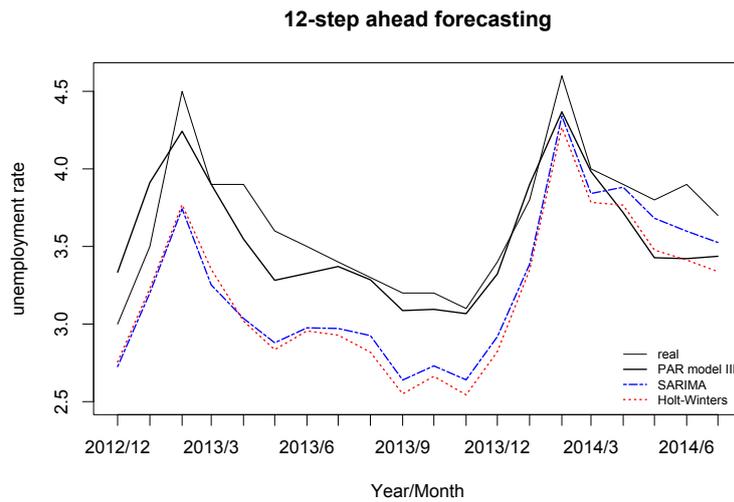


Figure 4.6. 12-step-ahead forecasting performances of the three models in the out-sample.

이상에서 볼 때, 다른 계절 모형에 비해 PAR 모형은 전반적으로 미래 예측시점에 관계없이 우수한 예측성능을 보이는 것을 확인하였다. 즉, 주기적 상관성을 지닌 계절형 시계열 자료 분석에 있어서 계절에 따라 모수값을 다르게 설정하는 PAR 모형이 예측에 있어 우수하며, 특히 모수 절약 모형의 사용이 예측 성능을 높이는 것을 확인할 수 있었다.

5. 결론

본 논문에서는 모형의 계수가 주기적으로 변하는 유연한 구조를 가지고 있는 주기적 자기회귀 모형을 소

개하고 모수 절약 구조를 반영하여 모형을 개선하는 방법을 제안하고 있다. 또한, 국내 실업률 자료를 이용하여 SARIMA 및 Holt-Winters 모형과의 예측 성능을 비교하였다. 실증분석 결과, 6개월 이상의 장기적 예측으로 갈수록 PAR 모형의 우수성을 확인할 수 있었으며, 특히 기존 PAR 모형보다 모수 절약 PAR 모형이 더욱 좋은 예측 성능을 나타냄을 확인할 수 있었다. 계절적 상관성을 지닌 자료에서의 PAR 모형의 우수성과 그 개선 가능성을 보여주는 연구로 평가하며 향후 분기별 또는 월별 자료 이외의 고주파수(high-frequency) 시계열 자료에 적용 가능한 형태로 변형할 수 있는 잠재력을 가지고 있다고 볼 수 있다.

References

- Boswijk, H. P., Franses, P. H., and Haldrup, N. (1997). Multiple unit roots in periodic autoregression, *Journal of Econometrics*, **80**, 167–193.
- Box, G. E. P. and Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*, Francisco Holden-Day.
- Franses, P. H. and McAleer, M. (1998). Cointegration analysis of seasonal time series, *Journal of Economic Survey*, **12**, 651–678.
- Franses, P. H. and Paap, R. (2004). *Periodic Time Series Models*, Oxford University Press.
- Holt, C. C. (1957). Forecasting trends and seasonals by exponentially weighted averages, ONR memorandum no. 52. Carnegie Institute of Technology, Pittsburgh, USA published in *International Journal of Forecasting* 2004, **20**, 5–10.
- Lee, S. D., Kim, J. G., and Kim, S. W. (2012). Estimation of layered periodic autoregressive moving average models, *Communications for Statistical Applications and Methods*, **19**, 507–516.
- Lund, R., Shao, Q., and Basawa, I. (2006). Parsimonious periodic time series modeling, *Australian & New Zealand Journal of Statistics*, **48**, 33–47.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer Science & Business Media.
- Matas-Mir, A. and Osborn, D. R. (2004). Does seasonality change over the business cycle? an investigation using monthly industrial product series, *European Economic Review*, **48**, 1309–1332.
- McLeod, A. I. (1993). Parsimony, model adequacy and periodic correlation in time series forecasting, *International Statistical Review/Revue Internationale de Statistique*, 387–393.
- Osborn, D. R. (1991). The implications of periodically varying coefficients for seasonal time-series processes, *Journal of Econometrics*, **48**, 373–384.
- Osborn, D. R. and Smith, J. P. (1989). The performance of periodic autoregressive models in forecasting seasonal U.K. consumption, *Journal of Business & Economic Statistics*, **7**, 117–127.
- Pagano, M. (1978). On periodic and multiple autoregressions, *The Annals of Statistics*, 1310–1317.
- Troutman, B. M. (1979). Some results in periodic autoregression, *Biometrika*, **66**, 219–228.
- Ursu, E. and Turkman, K. F. (2012). Periodic autoregressive model identification using genetic algorithms, *Journal of Time Series Analysis*, **33**, 398–405.
- Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages, *Management Science*, **6**, 324–342.

모수 절약 주기적 자기회귀 모형에 관한 연구

이지호^a · 성병찬^{a,1}

^a중앙대학교 응용통계학과

(2015년 12월 15일 접수, 2016년 1월 2일 수정, 2016년 1월 2일 채택)

요약

본 논문에서는 주기적 자기회귀(periodic autoregressive) 모형에서 모수의 수를 줄이기 위한 모수 절약 주기적 자기회귀 모형을 연구하였다. 제안된 모수 절약 모형은 실증분석에서 실업률을 이용하여 기존의 계절 시계열 모형과 비교를 통하여 그 성능을 평가하였다. 모수 절약 구조를 부여하기 위하여 계절성에서 공통된 패턴을 찾아내는 방법을 사용하였으며 기존 주기적 자기회귀 모형과의 통계적 차이 유무는 LR 검정을 통해 확인하였다. 그 결과, 중장기적으로 주기적 자기회귀 모형이 기존의 계절시계열 모형보다 우수한 예측성능을 보였으며, 특히 모수 절약 주기적 자기회귀 모형의 사용은 기존의 주기적 자기회귀 모형보다 우수한 예측성능을 나타내는 것을 확인하였다.

주요용어: 계절 시계열 모형, 계절성, ARIMA 모형, Holt-Winters 모형, 실업률

¹교신저자: (06974) 서울시 동작구 흑석로 84, 중앙대학교 경영경제대학 응용통계학과.
E-mail: bcseong@cau.ac.kr