

## Time Series Models for Daily Exchange Rate Data

Bomi Kim<sup>a</sup> · Jaehee Kim<sup>a,1</sup>

<sup>a</sup>Department of Statistics, Duksung Women's University

(Received July 2, 2012; Revised August 28, 2012; Accepted November 23, 2012)

---

### Abstract

ARIMA and ARIMA+IGARCH models are fitted and compared for daily Korean won/US dollar exchange rate data over 17 years. A linear structural change model and an autoregressive structural change model are fitted for multiple change-point estimation since there seems to be structural change with this data.

Keywords: ARIMA, BIC, Exchange rate, GARCH, IGARCH, structural change-points, structural change time series.

---

### 1. 서론

환율(exchange rate)이란 외국화폐로 표시한 자국화폐의 가격으로 각국 통화의 가치가 특정한 금속이나 다른 표준으로 측정되는 경우 환율은 '고정'된다. 그리고 수요·공급의 변동에 의해 환율이 결정되는 경우 환율은 '변동'한다. 만약 한 나라의 수입이 증가하면 재화의 수요가 증가하여 그 나라의 환율은 평가 절하될 것이며, 따라서 수입된 재화의 가격이 상승할 것이다. 재화의 가격이 상승함에 따라 수요는 감소하고 자국통화의 가치는 다른 나라의 통화에 비해 상대적으로 떨어지게 된다.

환율이 변동하는 데에는 여러 가지 경제적, 사회적인 요인들이 작용한다. 주요 요인들로는 금리의 변동과 경제성장률, 국제 수지 등이 있고 이외에도 정치적인 요인들도 환율변동의 주요한 원인들이 된다. 1997년 IMF 경제위기에 처했을 때와 2008년 전 세계적인 금융위기 때의 환율이 급등하면서 경제적으로 많은 어려움을 겪었다. 환율은 수출입에 가장 큰 영향을 주는 요소로 예상치 못한 환율 급등은 수출입을 하는 기업들에게 막대한 손해를 끼치게 된다. 이러한 손해를 예방하기 위하여 환율의 예측은 중요한 관심 문제이며 예측을 위한 연구가 활발히 진행되고 있다. 최근 조건부 분산이 변하는 경우의 경제 데이터에 대한 모형에 관심이 증가하고 있으며 이러한 상황을 반영하는 모형으로 GARCH 모형을 들 수 있다.

기존의 환율 분석 연구를 살펴보면 Jung (2005)은 개입모형과 지수평활법을 이용하여 환율의 최근 동향을 파악하고 이에 근거한 예측값을 산출하였다. Han (2003)은 원/달러 일일 환율과 외환 시장에서 한국 은행의 개입의 효과를 분석하는데 FIGARCH 모형으로 분석하였다. Lee (2006)은 1990년 1월부터 2005년 12월까지의 주별/월별 환율 데이터에 대해 화폐 환율변화율의 시계열에 조건부 변동성이 상

---

This research was supported by Basic Science Research Program through the National Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0011094).

<sup>1</sup>Corresponding author: Professor, Department of Statistics, Duksung Women's University, 33, Samyangro 144-gil, Dobong-gu, Seoul 132-714, Korea. E-mail: [jaehee@duksung.ac.kr](mailto:jaehee@duksung.ac.kr)

관관계를 갖는 ARCH 효과가 있음을 발견하고 적합한 모형으로 AR(1)+GARCH(1, 1)을 제안하였다. Moon (2010)은 환율 데이터에 대해 GARCH(1, 1) 모형을 적용하여 변동성을 추정하였다.

본 연구에서는 최근 17년간 환율 시계열 데이터에 대해 적합한 통계적 모형을 찾고 향후 환율을 예측해 보는 것에 의의를 두고 특히 ARIMA+IGARCH 모형과 구조변화모형 적합을 시도해 보고자 한다. 2장에서는 정상성 시계열 ARIMA 모형, 변동성 GARCH 모형과 구조 변화 모형에 대해 설명하고 3장에서 실제 환율 데이터에 대한 분석을 수행한 결과를 보여준다.

## 2. 시계열 데이터 분석 모형

### 2.1. ARIMA 모형

정상성(stationarity)은 시점에 관계없이 시계열이 일정한 특성을 가진다는 것을 의미한다. 정상성을 만족하려면 데이터의 평균이 일정하고, 분산이 시점( $t$ )에 의존하지 않으며, 공분산은 시차(lag)에만 의존해야 하고 이를 만족하지 않는 경우 비정상(nonstationary) 시계열이라고 한다. 시계열의 정상성을 검정하는 방법으로 널리 사용되는 augmented Dickey-Fuller(ADF)검정 (Dickey와 Fuller, 1981)이 있다.

$$H_0 : \text{단위근(unit root)이 존재한다}$$

에 대한 검정으로 귀무가설이 기각되면 시계열 데이터는 정상성을 만족한다고 볼 수 있다. 귀무가설을 기각하지 못하는 경우 비정상 시계열 데이터가 정상성을 만족하도록 차분(differentiation)을 통하여 정상 시계열로 바꾸어 줄 수 있다.

정상성을 만족하는 시계열에 대해서는 ARMA(Autoregressive moving average) 모형 또는 차분을 이용한 ARIMA(Autoregressive integrated moving average) 모형을 적합할 수 있다.

시계열 데이터  $\{Z_1, Z_2, \dots, Z_n\}$ 가 자기회귀과정(AR; autoregressive)과 이동평균과정(MA; moving average)을 혼합한 형태인 ARMA( $p, q$ ) 과정은 다음의 모형

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.1)$$

으로 표현한다. 여기서  $\{a_t\}$ 는 백색잡음과정(white noise process)으로 서로 독립이고 평균 0, 분산  $\sigma_a^2$ 인 확률 변수이다. ARMA( $p, q$ ) 과정에서 AR 차수  $p$ 의 결정에는 부분자기상관함수(PACF; partial autocorrelation function)를 MA 차수  $q$ 의 결정에는 자기상관함수(ACF; autocorrelation function)를 이용한다.

차분(difference)이란 현 시점의 데이터에서 전 시점의 데이터를 빼주는 것을 말한다. 평균이 일정하지 않은 비정상 시계열은 차분을 통하여 정상 시계열로 바꿀 수 있다.  $d$ 번 차분한  $\nabla^d Z_t \equiv W_t$ 가 ARMA( $p, q$ ) 모형을 따를 때  $Z_t$ 가 ARIMA( $p, d, q$ ) 모형을 따른다고 한다. 즉  $W_t$ 가 다음의 통계적 모형 ARMA( $p, q$ )

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.2)$$

를 만족한다.

### 2.2. GARCH와 IGARCH 변동성 모형

환율과 같은 금융시계열 데이터는 시간에 따른 분산의 변동이 크기 때문에 Box와 Jenkins (1976)가

제안한 선형시계열 모형인 ARIMA 모형으로는 설명하기 어려운 경우가 많다. 이러한 변동성을 설명하기 위하여 Engle (1982)은 ARCH(autoregressive conditional heteroscedasticity) 모형을 제안하였으며 Bollerslev (1986)는 일반화된 ARCH 모형으로 GARCH(generalized autoregressive conditional heteroscedasticity) 모형을 제안하였다. Tsay (2005)는 이들 모형에 대한 자세한 설명을 제공한다.

변동성을 갖는 확률변수  $\{a_t\}$ 에 대해 GARCH( $p, q$ ) 모형은 다음과 같다:

$$\begin{aligned} a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i a_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \end{aligned} \quad (2.3)$$

여기서  $\omega > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$ 을 만족하며  $\{\epsilon_t\}$ 는 평균 0, 분산 1을 갖는 서로 독립이고 동일한 분산을 갖는 확률 변수이다.

$a_t$ 의 기대값과 분산을 계산해보면

$$E(a_t) = E(\sigma_t)E(\epsilon_t) = 0, \quad (2.4)$$

$$\begin{aligned} \text{Var}(a_t) &= E(a_t^2) = E(\sigma_t^2) = \omega + \sum_{i=1}^q \alpha_i E(a_t^2) + \sum_{j=1}^p \beta_j E(\sigma_t^2) \\ &= \omega + \left( \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right) E(a_t^2) \end{aligned} \quad (2.5)$$

이므로 분산은

$$E(a_t^2) = \frac{\omega}{1 - \left( \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right)} \quad (2.6)$$

이다.

GARCH( $p, q$ ) 모형이 유일한 정상 과정(stationary process)이기 위한 필요충분조건은

$$E(a_t^2) < \infty, \quad \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1 \quad (2.7)$$

이다 (Theorem 4.4, Fan과 Yao, 2002).

GARCH 모형은 현재의 변동성이 미래에는 지수적으로 감소하여 없어지는 모형이다. Nelson (1990)은 현재의 변동성이 미래의 변동성에 그대로 남아있는 경우에 대한 모형으로 Integrated-GARCH(IGARCH) 모형을 제안하였다.

IGARCH(1,1) 모형식은 다음과 같다:

$$\begin{aligned} a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \omega + \alpha a_{t-1}^2 + (1 - \alpha) \sigma_{t-1}^2, \end{aligned} \quad (2.8)$$

여기서  $\omega \geq 0$  그리고  $\alpha \geq 0$ 을 만족하며  $\{\epsilon_t\}$ 는 평균 0, 분산 1을 갖는 서로 독립이고 동일한 분산을 갖는 확률 변수이다.

### 2.3. 모형 평가

모형을 적합한 후에는 모형에 대한 평가를 해야한다. 이러한 방법으로 잔차에 대한 독립성 검정과 정보 기준통계량을 이용한 모형 비교에 대해 설명하고자한다.

적절한 모형이 적합된 후 잔차들은 독립성을 만족해야하는데 ARMA( $p, q$ ), GARCH( $p, q$ ) 모형에 대해 잔차의 자기상관성 검정으로 Ljung과 Box (1978) 가 제안한 Ljung-Box 검정을 할 수 있다. 즉  $H_0$  : 데이터가 독립이다 라는 귀무가설에 대해 Ljung-Box 검정통계량은

$$Q(g) = n(n+2) \sum_{k=1}^g \frac{\hat{\rho}_k^2}{n-k} \quad (2.9)$$

이다. 여기서  $n$ 은 표본크기,  $\hat{\rho}_k$ 은 lag  $k$ 에서 표본자기상관이고  $g$ 는 검정에 사용하는 lag값이다. 유의수준  $\alpha$ 에서 검정법은  $Q(g) > \chi_{1-\alpha, g}^2$ 이고 여기서  $\chi_{1-\alpha, g}^2$ 은 자유도가  $g$ 인 카이제곱 분포의 오른쪽 꼬리로부터  $\alpha$ -분위수이다.

데이터에 대한 정규성 검정으로 왜도와 첨도를 이용한 Jarque-Bera 검정을 할 수 있다. ' $H_0$  : 왜도와 첨도가 정규분포의 왜도와 첨도와 같다'라는 귀무가설에 대해 Jarque-Bera 검정통계량

$$JB = \frac{n}{6} \left[ S_w^2 + \frac{(K_u - 3)^2}{4} \right] \quad (2.10)$$

은  $H_0$ 하에서 자유도 2인 카이제곱분포를 따른다. 여기서  $S_w$ 는 잔차에 대한 왜도(skewness)이고  $K_u$ 는 첨도(kurtosis)이다.

정보기준통계량을 통해 시계열 모형을 평가하는 방법으로는 AIC(Akaike Information Criterion)와 BIC(Bayesian Information Criterion)이 널리 쓰인다. AIC 또는 BIC 값을 최소로 하는 모형을 적절한 모형으로 선택할 수 있으며 여러 가지 모형을 비교할 때 유용하다. 각 통계량은

$$AIC = -2 \log(\text{maximum likelihood}) + 2K, \quad (2.11)$$

$$BIC = -2 \log(\text{maximum likelihood}) + K \log(n) \quad (2.12)$$

이다. 여기서  $K$ 는 모형에서 모수의 개수이고  $n$ 은 표본크기이다.

### 2.4. 예측

시계열 데이터를 모형화 하는 중요한 목적은 분석한 모형을 이용하여 미래에 대한 예측(forecasting)이다. 예측이란  $t$ 시점까지의 시계열 데이터  $Z_1, Z_2, \dots, Z_t$ 를 가지고  $k$ 시점 앞의 값  $Z_{t+k}$ 를 예측하는 것으로 시점  $t$ 에서 선시차가  $k$ 인 시계열  $Z_1, Z_2, \dots, Z_t$ 의 최소평균제곱오차 예측은

$$\hat{Z}_t(k) = E(Z_{t+k} | Z_t, Z_{t-1}, \dots, Z_1) \quad (2.13)$$

으로 데이터  $Z_1, Z_2, \dots, Z_t$ 가 주어졌을 때  $k$ 시점 앞의 시계열 예측값은  $Z_{t+k}$ 의 조건부 기대값으로 나타난다. 예측오차(forecast error)는

$$e_t(k) = Z_{t+k} - \hat{Z}_t(k) \quad (2.14)$$

이고 기대값  $E[e_t(k)] = 0$ 이다. 분산  $\text{Var}[e_t(k)]$ 을 구하여 근사적인 95% 예측신뢰구간을

$$\hat{Z}_t(k) \pm 2\sqrt{\{\text{Var}[e_t(k)]\}} \quad (2.15)$$

으로 구한다.

## 2.5. 구조 변화 시계열 모형

특정한 정책의 시행, 국내외적인 경제적 사건 등으로 경제 시계열 데이터 구조의 변화가 일어날 수 있으며 구조 변화는 시간에 따라 시계열의 모수들이 변함을 의미한다.

시계열  $\{Z_t\}$ 에 대해 선형구조를 포함한 구조 변화 모형 (i)과 시계열 모형을 포함한 구조 변화 모형 (ii)를 고려하고자 한다.  $j = 1, 2, \dots, m+1$ 에 대하여

$$(i) \quad Z_t = \beta_{j0} + \beta_{j1}t + \epsilon_t, \quad t = c_{j-1} + 1, \dots, c_j, \quad (2.16)$$

$$(ii) \quad Z_t = \beta_{j0} + \beta_{j1}t + \phi_{1j}Z_{t-1} + \phi_{2j}Z_{t-2} + \epsilon_t, \quad t = c_{j-1} + 1, \dots, c_j, \quad (2.17)$$

여기서 변화점 개수  $m$ 은 정해지지 않았으며  $m$ 개 변화점  $c_1, \dots, c_m$ 에 대해  $1 = c_0 < c_1 < c_2 < \dots < c_m < c_{m+1} = n$ 이 성립한다. 즉  $m$ 개 변화점에 의해  $m+1$ 개 부집단(segment)이 형성된다.

어떤 모형의 구조적 변화에 대한 증거가 있다고 볼 수 있다면, 자연스러운 전략은 변화를 포함한 모형을 찾는 것이다. 갑작스러운 변화(abrupt change)가 존재할 때 변화점에 의해 데이터를 부집단으로 나눌 수 있으며 변화점으로 부집단이 분리된 후에는 각 부집단별 모형을 적합하고 모수를 추정해야한다. Bai와 Perron (1998, 2003)은 이러한 절단점(break point) 또는 다중 변화점(change-point)을 추정하는 방법을 제안하였다. 이들의 변화점 추정 방법은 Zeileis 등 (2003)가 R 패키지로 만들어 R 함수를 제공한다.

Bai와 Perron (2003)의 다중변화점 추정방법은  $m+1$ 개 부집단에 대해 잔차제곱합(residual sum of squares; RSS)을 최소화하는 변화점을 찾기위해 dynamic programming 알고리즘을 사용한다. 여기서  $m$ 은 변화점 개수로 모르는 값이다. 띠틈(bandwidth)  $h$ ,  $0 < h < 1$ 를 정해 각 부집단에는 최소  $h \cdot n$ 개의 관측값들이 포함되도록 하였다. 각 부분에서 관측값들의 최소한의 비율이 일반적으로 10% 또는 15%로 선택된다. 또한 각 모형에 대해 BIC를 계산하여 BIC를 최소화하는 모형을 선택한다.

$$(\hat{c}_1, \dots, \hat{c}_m) = \arg \min_{c_1, \dots, c_m} S_n(c_1, \dots, c_m), \quad (2.18)$$

여기서 선형구조를 포함한 구조 변화 모형 (i)과 시계열 모형을 포함한 구조 변화 모형 (ii)에 대한 모형의 잔차제곱합은 각각

$$(i) \quad S_n(c_1, \dots, c_m) = \sum_{j=1}^m \sum_{t=c_{j-1}}^{c_j} (Z_t - \beta_{j0} - \beta_{j1}t)^2 \quad (2.19)$$

$$(ii) \quad S_n(c_1, \dots, c_m) = \sum_{j=1}^m \sum_{t=c_{j-1}}^{c_j} (Z_t - \beta_{j0} - \beta_{j1}t - \phi_{1j}Z_{t-1} - \phi_{2j}Z_{t-2})^2 \quad (2.20)$$

이다.  $c_i - c_{i-1} \geq q = h \cdot n$ 인 모든 가능한 partition에 대해 잔차제곱합을 최소화하는 다중 변화점을 추정한다.

## 3. 환율 데이터에 대한 시계열 분석

### 3.1. 환율 데이터

본 연구에서 사용한 환율 원/달러 데이터는 1995년 1월 3일부터 2012년 1월 31일까지 17년간 일별 환율 데이터로 총 데이터 개수는 4615개 이며 외환은행 환율고시의 매매기준율(basic rate of exchange)을 사용하였고 단위는 원/달러이다. 환율 데이터는 Figure 3.1에서 보여주며 1997년 한국의 외

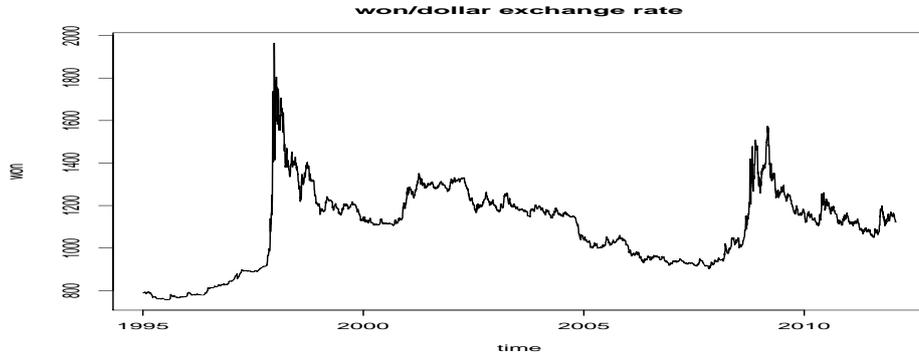


Figure 3.1. Won/dollar exchange rate time series data

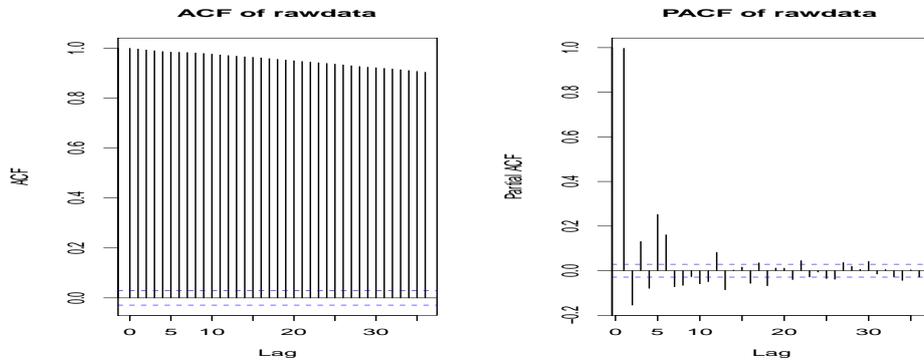


Figure 3.2. (a) ACF and (b) PACF of won/dollar exchange rate time series data

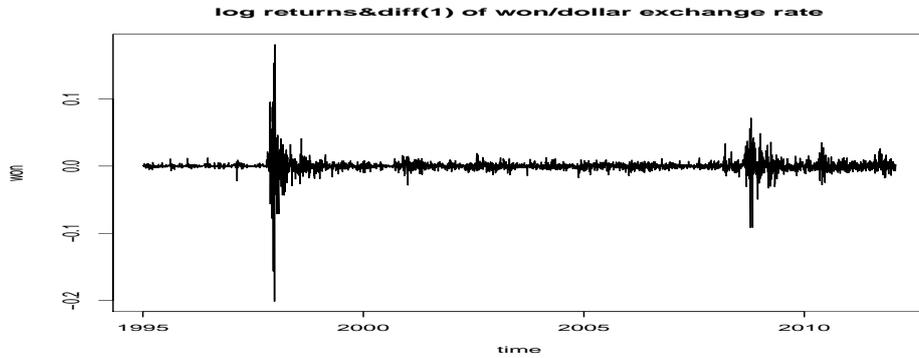


Figure 3.3. Log transformation and first differenced exchange rate data

환위기로 환율이 급등 했었고 2001년 미국의 9.11테러와 2009년 서브프라임 금융위기로 환율이 증가하는 경향이 보인다. 데이터가 증가와 감소를 반복하고 있으며 변동 폭이 일정하지 않으므로 비정상 시계열 데이터로 의심이 된다. 본 연구에서는 이러한 환율 데이터에 대한 적합한 시계열 모형을 찾고자 하며 R 프로그램에서 `tseries`, `forecast`, `TSA`, `fGarch`, `rugarch`와 `strucchange` 패키지를 활용하였다.

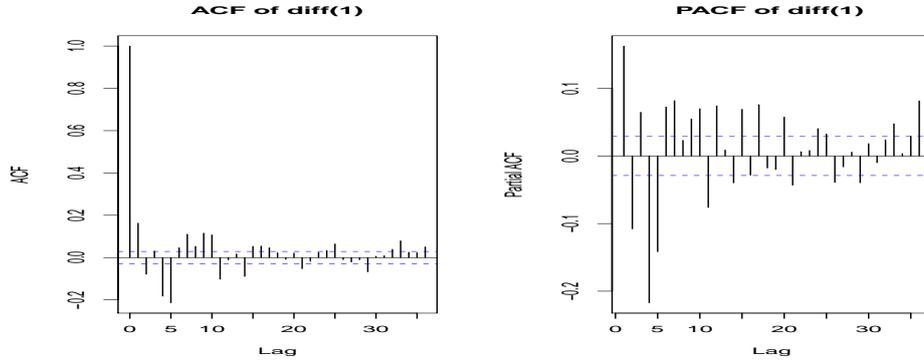


Figure 3.4. (a) ACF and (b) PACF of log transformation and first differenced exchange rate data

Table 3.1. Augmented dickey-fuller(ADF) test

	dickey-fuller statistics	lag	$p$ -value
raw data	-2.5129	16	0.3612
log transformation	-2.2302	16	0.4809
first differenced	-14.437	16	0.01
log transformation and first differenced	-14.071	16	0.01

시계열 데이터의 정상성 여부를 확인하기 위해서 자기상관함수(ACF)와 부분 자기상관함수(PACF) 그림을 그려보았다. Figure 3.2(a) 그래프를 보면 ACF 값이 서서히 줄어들지만 그 값이 1에 가까워 매우 크다. Figure 3.2(b)에서 PACF 그래프를 보면 lag = 1에서 값이 1에 가깝고 이후에 급격히 줄어들었음을 확인 할 수 있다. 그러므로 자료가 비정상 시계열로 판단된다. 그러므로 로그변환한 후에 1차 차분한 데이터  $\{W_t\}$ 에 대한 시계열 그래프 Figure 3.3을 보면 로그변환을 통하여 분산은 대체로 안정화 되고 1차 차분을 통하여 평균이 일정해 진 것을 볼 수 있다. 시계열의 정상성 여부를 확인하기 위해서  $\{W_t\}$ 에 대해 자기상관함수와 부분자기상관함수를 그려보면 Figure 3.4(a) ACF, (b) PACF와 같다.

정상성 판단을 위하여 Augmented Dickey-Fuller(ADF) 검정을 실시한 결과 Table 3.1에서 1차 차분한 데이터와 로그변환과 1차 차분한 데이터의  $p$ -값 = 0.01이므로 정상성을 만족한다고 보아 로그변환과 1차 차분한 데이터를 선택하여 분석하기로 한다.

### 3.2. ARIMA 모형 적합

원 데이터를 로그변환과 1차 차분한 데이터  $\{W_t\}$ 에 대해 ARIMA 모형을 적합하고자 한다. Table 3.2에서 ARIMA( $p, 1, q$ ) 모형의 오차 분산 추정값과 로그우도, AIC, BIC 등을 비교한 결과 AIRMA(2, 1, 2)의 모형을 선택하기로 한다. 적합한 모형의 식은 다음과 같다.

$$W_t = 0.3599W_{t-1} + 0.2131W_{t-2} + a_t - 0.1725a_{t-1} - 0.4193a_{t-2}, \quad (3.1)$$

여기서  $t = 1, 2, \dots, n-1$ 이고  $a_t$ 는 백색잡음 과정으로 서로 독립이고 평균 0, 분산  $\sigma_a^2$ 인 확률변수이다.

### 3.3. ARIMA+GARCH 모형 적합

로그변환과 1차 차분한 환율 데이터  $\{W_t\}$ 에서 시간에 따라 변동성이 보이므로 변동성에 대해 GARCH 모형을 고려해 보자. Figure 3.4를 보면  $\{W_t\}$ 에 대해 자기회귀상관성이 보이므로 ARIMA+GARCH

**Table 3.2.** Comparison of log transformation ARIMA( $p, 1, q$ ) models

model	$\hat{\sigma}^2$	likelihood	AIC	BIC
ARIMA(2, 1, 3)	$8.214 \times 10^{-5}$	15155.12	-30300.24	-30261.62
ARIMA(2, 1, 2)	$8.216 \times 10^{-5}$	15154.51	-30301.02	-30268.83
ARIMA(2, 1, 1)	$8.251 \times 10^{-5}$	15144.84	-30283.69	-30257.94
ARIMA(2, 1, 0)	$8.348 \times 10^{-5}$	15117.72	-30231.44	-30212.13
ARIMA(1, 1, 0)	$8.447 \times 10^{-5}$	15090.75	-30179.50	-30166.62
ARIMA(1, 1, 1)	$8.257 \times 10^{-5}$	15143.08	-30282.16	-30262.85
ARIMA(1, 1, 2)	$8.252 \times 10^{-5}$	15144.50	-30283.00	-30257.25
ARIMA(0, 1, 1)	$8.383 \times 10^{-5}$	15108.05	-30214.11	-30201.23
ARIMA(0, 1, 2)	$8.267 \times 10^{-5}$	15140.38	-30276.76	-30257.45

**Table 3.3.** Comparison of ARIMA( $p, d, q$ )+GARCH(1, 1) models

model	information criterion		Ljung-Box test		Jarque-Bera test	
	AIC	BIC	Q(10) for $a_t^2$	$p$ -value	statistic	$p$ -value
ARIMA(1, 0, 0)+GARCH(1, 1)	-8.1533	-8.1478	10.42496	0.4040	3790.45	0
ARIMA(2, 0, 0)+GARCH(1, 1)	-8.1699	-8.1643	9.433257	0.4915	3614.62	0
ARIMA(1, 1, 0)+GARCH(1, 1)	-8.1743	-8.1688	9.460000	0.4891	3533.75	0
ARIMA(2, 1, 0)+GARCH(1, 1)	-8.1662	-8.1607	10.24231	0.4195	3347.29	0
ARIMA(1, 1, 1)+GARCH(1, 1)	-8.1582	-8.1527	8.879696	0.5436	3442.35	0

**Table 3.4.** Comparison of ARIMA( $p, d, q$ )+IGARCH(1, 1) models

model	information criterion		Ljung-Box test	
	AIC	BIC	Q(10) for $a_t^2$	$p$ -value
ARIMA(1, 0, 0)+IGARCH(1, 1)	-8.1532	-8.1462	10.43	0.3171
ARIMA(0, 0, 1)+IGARCH(1, 1)	-3.2685	-3.2615	1557	0
ARIMA(1, 0, 1)+IGARCH(1, 1)	-2.1381	-2.1325	160.7	0
ARIMA(1, 1, 0)+IGARCH(1, 1)	-8.1704	-8.1634	10.86	0.2857
ARIMA(0, 1, 1)+IGARCH(1, 1)	-8.1702	-8.1632	10.67	0.2993
ARIMA(1, 1, 1)+IGARCH(1, 1)	-8.1521	-8.1465	11.66	0.3088

모형 적합을 시도하고자 한다. 로그변환 데이터  $\{Y_t\}$ 에 대해 ARIMA( $p, d, q$ )+GARCH(1, 1) 모형에 대한 모형적합도 통계량들은 Table 3.3에서 보여준다. AIC와 BIC 값을 최소로 하고 Ljung-Box 검정을 이용하여 데이터가 독립이라는 귀무가설을 채택하는 모형인 ARIMA(1, 1, 0)+GARCH(1, 1) 모형을 최종적으로 선택하여 다음의 모형

$$\begin{aligned}
 W_t &= 0.1627W_{t-1} + a_t, \quad t = 2, \dots, n, \\
 a_t &= \sigma_t \epsilon_t, \\
 \sigma_t^2 &= 1.943 \times 10^{-7} + 0.1957a_{t-1}^2 + 0.8270\sigma_{t-1}^2
 \end{aligned} \tag{3.2}$$

으로 구하였다. 여기서  $\epsilon_t$ 는 서로 독립이고 평균 0, 분산 1인 확률변수이다.

Jarque-Bera 검정결과  $p$ -값이 매우 작게 나와 정규분포의 특성을 만족하지 않는 것으로 여겨진다. 모형 적합식에서 모수의 합이  $\alpha + \beta > 1$ 로 정상성을 만족하지 못한다. 환율 데이터인 경우에 모수의 합이 1보다 큰 경우가 종종 있으며 이러한 경우  $\alpha + \beta = 1$ 인 조건을 주는 IGARCH 모형적합을 시도할 수 있다.

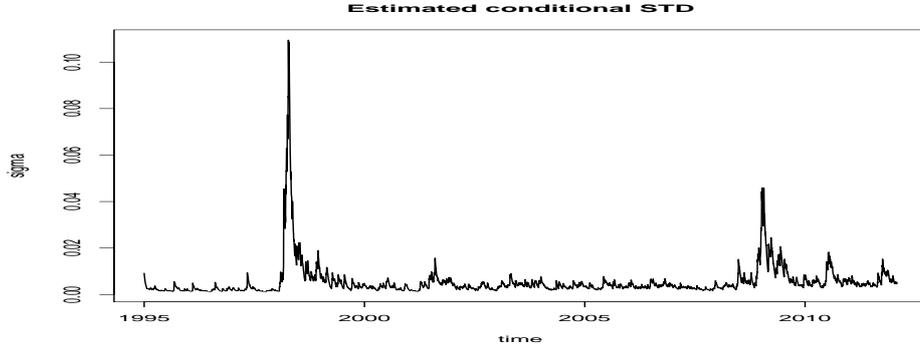


Figure 3.5. Conditional standard deviation estimates of ARIMA+IGARCH model

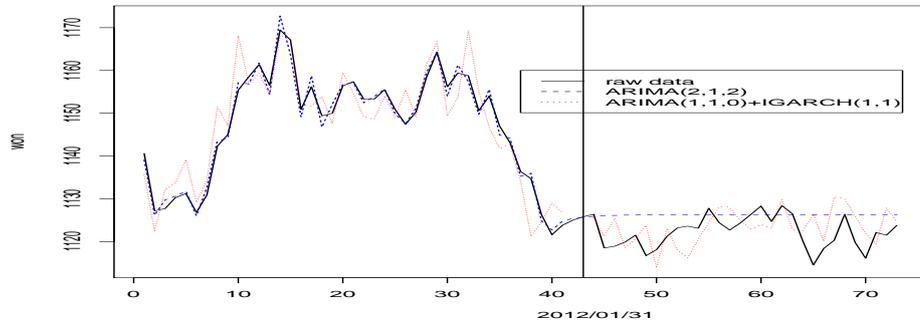


Figure 3.6. Estimates and prediction of ARIMA model and ARIMA+IGARCH model

### 3.4. ARIMA+IGARCH 모형 적합

ARIMA(1, 1, 0)+GARCH(1, 1) 모형의 모수의 합이 1이 넘어 정상성을 만족시키지 못하여 IGARCH 모형을 적합해 보았다. Table 3.4에서 ARIMA( $p, d, q$ )+IGARCH(1, 1) 모형들에 대한 모형 적합도 통계량들을 보여주고 이 중에서 AIC와 BIC 값이 가장 작고 Ljung-Box 검정결과를 만족시키는 ARIMA(1, 1, 0)+IGARCH(1, 1) 모형을 최종적으로 선택하고 추정된 모형은 다음과 같다.

$$\begin{aligned} W_t &= 0.1627W_{t-1} + a_t, \quad t = 2, \dots, n, \\ a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= 0.1616a_{t-1}^2 + 0.8383\sigma_{t-1}^2, \end{aligned} \quad (3.3)$$

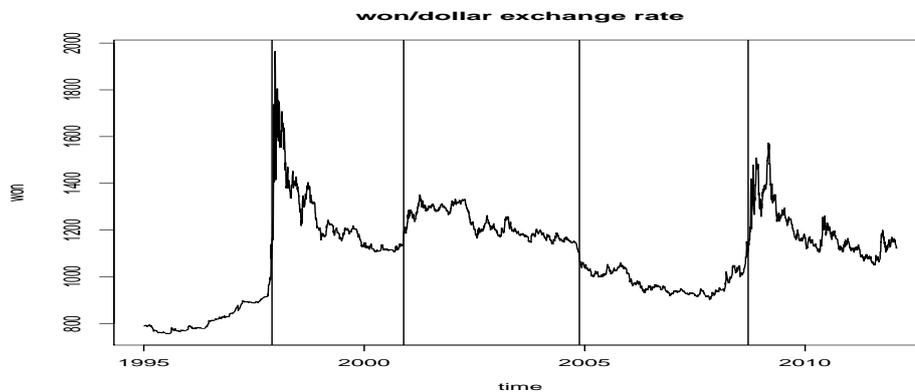
여기서  $\epsilon_t$ 는 서로 독립이고 평균 0, 분산 1인 확률변수이다. ARIMA(1, 1, 0)+IGARCH(1, 1) 모형에 의해서 추정된 조건부표준편차(conditional standard deviation)  $\hat{\sigma}_t$ 는 Figure 3.5에서 보여준다.

### 3.5. ARIMA(2, 1, 2) 모형과 ARIMA(1, 1, 0)+IGARCH(1, 1) 모형의 비교

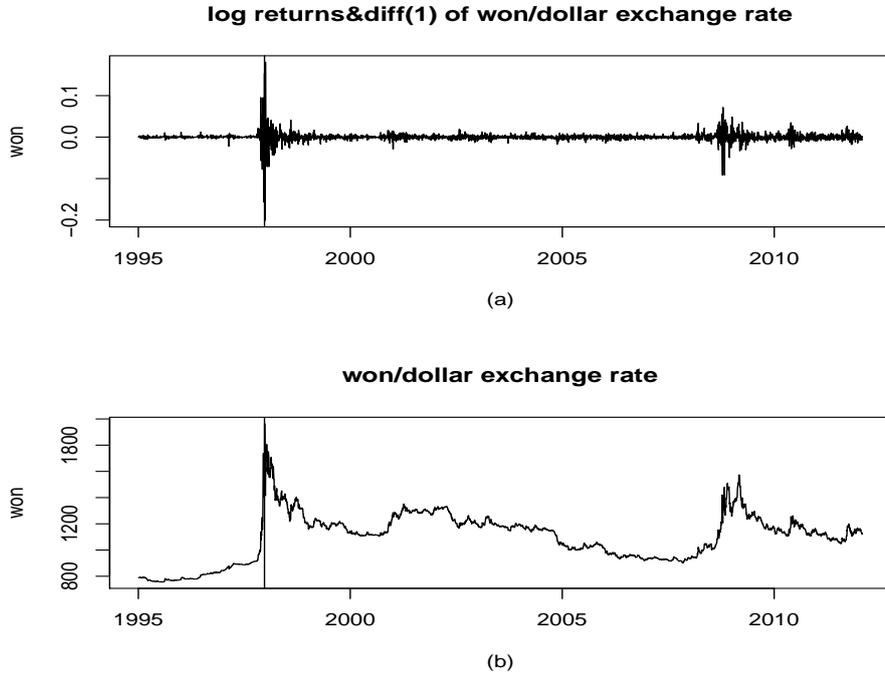
최종 선택한 ARIMA(2, 1, 2) 모형과 ARIMA(1, 1, 0)+IGARCH(1, 1) 모형에 대해 비교해보고자 한다. 1995년 1월부터 2012년 1월 31일까지 데이터에 대해 두 모형을 적합하였으며 향후 한달간의 환율에 대한 예측을 실시 하였다. Figure 3.6은 모형적합에 사용된 마지막 2개월간 두 모형의 추정값과 향후 한달

**Table 3.5.** Comparison of ARIMA(2, 1, 2) and ARIMA(1, 1, 0)+IGARCH(1, 1) models

date	real data	ARIMA prediction	ARIMA+IGARCH prediction
2012-02-01	1125.8	1125.404	1125.601
2012-02-02	1126.4	1125.702	1121.244
2012-02-03	1118.5	1125.896	1125.372
2012-02-06	1118.9	1126.029	1118.578
2012-02-07	1119.9	1126.118	1120.633
2012-02-08	1121.5	1126.179	1123.742
2012-02-09	1116.7	1126.219	1113.906
2012-02-10	1118.1	1126.247	1123.033
2012-02-13	1121.2	1126.266	1117.994
2012-02-14	1123.2	1126.278	1116.221
2012-02-15	1123.6	1126.287	1120.559
2012-02-16	1123.1	1126.292	1124.427
2012-02-17	1127.8	1126.296	1128.119
2012-02-20	1124.4	1126.299	1128.034
2012-02-21	1122.7	1126.301	1124.879
2012-02-22	1124.3	1126.302	1122.902
2012-02-23	1126.3	1126.303	1123.953
2012-02-24	1128.3	1126.303	1123.077
2012-02-27	1124.7	1126.304	1129.770
2012-02-28	1128.4	1126.304	1122.790
2012-02-29	1126.5	1126.304	1122.396

**Figure 3.7.** Change-point estimates with linear structural change model

동안의 환율 예측값을 비교한 그림이다. 과거 데이터에 대한 추정을 보면 환율 추세에 따라 대체적으로 잘 추정한 것으로 보인다. ARIMA(2, 1, 2) 모형의 예측값은 증가추세를 나타내고 약 한달 이후부터 일정한 것에 비해 ARIMA(1, 1, 0)+IGARCH(1, 1) 모형은 변동성을 잘 예측한 것으로 보인다. 두 추정값 모두 실제값과 차이가 큰 편은 아니다. 모형 적합에 사용한 과거의 데이터에 대한 추정보다 미래 환율 예측에서는 실제값과 좀 더 차이가 남을 볼 수 있다. 경제적 상황 또는 사건이 환율 변동에 영향을 미치는 경우에는 추정한 모형만으로는 예측에는 한계가 있기 때문이기도 하다. Table 3.5는 모형 적합 후 미래에 대한 환율 실제값과 두 모형에 의한 각 예측값을 보여주어 실제값과 예측값의 차이를 알 수 있다.



**Figure 3.8.** (a) Log transformation and first differenced exchange rate data, (b) Change-point estimates in structural change model with autocorrelation

### 3.6. 구조 변화 모형

환율에 대한 시계열 데이터의 구조가 전체 표본 기간 내내 동일하지 않게 보이므로 Bai와 Perron (2003) 방법으로 구조 변화점 추정을 해보고자 한다.

시계열  $\{Z_t\}$ 에 대해 선형구조를 포함한 구조 변화 모형과 자기상관 구조를 포함한 구조 변화 모형을 고려하고자 한다. (i) 모형에 대한 다중 변화점은 결과 862, 1748, 2815와 3766번째로 추정되었고 Figure 3.7에서 변화점의 위치를 보여준다. 추정된 변화점은 각각 1997년 11월 28일, 2000년 11월 23일, 2004년 11월 18일 그리고 2008년 9월 16일로 1997년 IMF 외환위기로 환율이 급등하여 최고값을 나타내고 2008년 리먼 브라더스의 파산으로 미국 금융위기가 발생하여 다시 환율이 급등했음을 반영한다. 변화점으로 분리되는 각 부분에 대해 모수를 추정하여 다음의 모형

$$Z_t = \begin{cases} 6.599614 + 0.000257t, & t = 1, \dots, 862, \\ 7.616580 - 0.000379t, & t = 863, \dots, 1748, \\ 7.411548 - 0.000131t, & t = 1749, \dots, 2815, \\ 7.101 - 6.565 \times 10^{-5}t, & t = 2816, \dots, 3766, \\ 8.192977 - 0.000265t, & t = 3767, \dots, 4615 \end{cases}$$

을 얻었다.

로그변환과 차분한 시계열  $\{W_t\}$ 인 Figure 3.3에 대해 자기상관을 포함한 모함한 구조 변화 모형 (ii)에 대한 변화점은 882번째로 추정되었고 이는 1997년 12월 23일로 Figure 3.8(a)와 (b)에서 변화점의 위

치를 보여준다. 변화점에 따른 구조 변화 모형은 다음과 같다:

$$W_t = \begin{cases} -2.1862 + 0.008318t + 0.3456W_{t-1}, & t \leq 882, \\ -0.1868 + 0.18750t + 0.0038W_{t-1}, & t \geq 883. \end{cases}$$

자기상관 구조를 포함한 차분 데이터에 대한 구조변화 모형은 로그변환과 1차 차분한 데이터를 이용한 것으로 원데이터와 비교하였을 때 2008년경의 환율 증가부분이 1997년때의 변동에 비하여 변동폭이 작음을 확인 할 수 있고 이로 인해 변화점으로 추정되지 않은것으로 보인다.

#### 4. 결론

미국 달러에 대한 한국원화의 17년간 일별 원/달러 환율 데이터에 대해 ARIMA 모형과 ARIMA+IGARCH 모형을 적합하여 비교하고 예측을 하였다. 시계열 데이터를 정상화 시키기 위하여 로그변환과 1차 차분을 시도하였고 ARIMA(2, 1, 2) 모형과 변동성을 포함한 ARIMA(1, 1, 0)+IGARCH(1, 1) 모형을 선택하여 적합하였다. 두 모형을 이용하여 추정해 본 결과 과거 데이터에 대한 추정을 잘 하는 편이고 미래 예측 결과는 ARIMA(2, 1, 2) 모형에 비해 ARIMA(1, 1, 0)+IGARCH(1, 1) 모형이 실제로 나타난 환율의 변동성을 잘 반영했다.

또한 환율 데이터에 구조 변화가 있어 보여 선형모형 구조 변화점 추정과 자기상관추세를 포함한 구조 변화점 추정을 하였다. 각 모형에 대한 변화점 추정 결과를 보면 변화점 개수와 변화점의 위치는 약간 다르게 나타났으나 1997년 IMF 외환위기로 인한 환율 급등은 두 모형에서 모두 변화점으로 추정되었다. 이는 특정한 경제적 사건이 환율 데이터의 변화점을 만든다고 할 수 있다.

시계열이면서 동적인(dynamic) 데이터는 금융/경제 분야에서 많이 발생하고 있으며 이러한 시계열 데이터에 대한 통계적 모형 연구는 금융통계/경제통계 분야에서 활발히 진행되고 있다. 앞으로 새로운 통계적 모형 창출과 데이터 상황을 최대한으로 반영한 창의적인 통계 분석이 기대된다.

#### References

- Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes, *Econometrica*, **66**, 47-78.
- Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models, *Journal of Applied Econometrics*, **18**, 1-22.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307-327.
- Box, G. E. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*, Francisco Holden-Day, New York.
- Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, **49**, 1057-1072.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987-1007.
- Fan, J. and Yao, Q. (2002). *Nonlinear Time Series*, Wiley, New York.
- Han, Y. W. (2003). Long memory property and central bank intervention during the currency crisis in the daily Korean won-us dollar exchange rates, *The Journal of the Korean Economy*, **4**, 93-116.
- Jung, D. B. (2005). Detection and forecasting of exchange rate using time series analysis, *Journal of Management Education*, 143-151.
- Lee, S. Y. (2006). A study on the behavior of foreign exchange rate volatility, *Korean Academy of International Business Management*, **2006.11**, 407-424.

- Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models, *Biometrika*, **65**, 553–564.
- Moon, C. K. (2010). Exchange rate volatility measures and GARCH model applications: Practical information processing approach, *International Commerce and Information Review*, **12**, 99–121.
- Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1,1) model, *Econometric Theory*, **6**, 318–334.
- Tsay, R. S. (2005). *Analysis of Financial Time Series*, Wiley-interscience, New York.
- Zeileis, A., Kleiber, C., Krämer, W. and Hornik, K. (2003). Testing and dating of structural changes in practice, *Computational Statistics & Data Analysis*, **44**, 109–123.

# 일별 환율데이터에 대한 시계열 모형 적합 및 비교분석

김보미<sup>a</sup> · 김재희<sup>a,1</sup>

<sup>a</sup>덕성여자대학교 정보통계학과

(2012년 7월 2일 접수, 2012년 8월 28일 수정, 2012년 11월 23일 채택)

---

## 요약

미국 달러에 대한 한국원화의 17년간 일별 원/달러 환율 시계열 데이터에 대하여 정상 시계열 ARIMA 모형과 변동성을 포함한 시계열 모형인 ARIMA+IGARCH 모형을 적합하여 비교하고 예측을 실시하였다. 또한 환율 데이터에 구조변화가 있어 보이므로 선형구조를 포함한 구조 변화 모형과 자기상관 구조를 포함한 구조 변화 모형을 이용하여 변화점을 추정하고자 한다.

주요용어: ARIMA 모형, BIC, GARCH 모형, IGARCH 모형, 구조변화 시계열 모형, 구조 변화점, 환율.

---

---

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No. 2010-0011094)입니다.

<sup>1</sup>교신저자: (132-714) 서울시 도봉구 삼양로 144길 33, 덕성여자대학교 정보통계학과, 교수.

E-mail: jaehee@duksung.ac.kr