Journal of Communication and Computer 12 (2015) 79-84
doi: 10.17265/1548-7709/2015.02.006

Nonlinear Time-varying AR-ARCH Model Based on Chaos Prediction Model and its Statistical Significance Tests

Tomoya Suzuki and Hajime Onuma

Department of Intelligent Systems Engineering, College of Engineering, Ibaraki University,4-12-1 Nakanarusawa-cho, Hitachi, Ibaraki 316-8511, Japan

Abstract: We proposed some nonlinear versions of the AR-ARCH model, which is often used for financial computing but is based on the linear regression. Our nonlinear AR-ARCH model can temporally change its model parameters following the local structure hidden in the observed financial data. To confirm the validity of our model, we performed some statistical significance tests. First, as one of the learning process, we performed the surrogate data test to the learning data, but we could not aggressively suggest that the original financial data have nonlinearity. However, from the viewpoint of predictive performance of the test data, our nonlinear models, especially the NAR-NARCH model, can improve the prediction accuracy of both of the return rate and the volatility. This improvement can be verified by the Wilcoxon signed-rank test.

Keywords: Financial computing, AR-ARCH model, nonlinear prediction model, surrogate data test.

1. Introduction

In the field of financial engineering and computing, the return rate and its standard deviation (called as volatility) are mainly used for asset management such as portfolio theory and option pricing theory. Therefore, some time-series models have been proposed to learn their behavioral patterns from the past historical movements and to estimate their possible future values. As typical examples, the AR model is used for the return rate, and the ARCH model [1] is used for the volatility. Then, these can be mixed together as the AR-ARCH model [2, 3]. All of the above models are categorized into linear models because they are based on the multivariate linear regression. However, according to some previous studies [4, 5], financial markets might be so complex that linear models are not good enough to express them, and therefore it would be possible that higher order nonlinear models could work better. From this viewpoint, the present study modifies the original AR-ARCH model into its nonlinear versions by applying a local linear approximation method [6], which is one of nonlinear predictions based on the chaotic dynamical theory. By this modification, the regression coefficients of the AR-ARCH model can be time-varying by approximating each local structure of the nonlinear dynamics hidden in financial time-series data.

Next, to confirm the validity of our modification, we perform two kinds of statistical significance tests: the surrogate data test [7] and the Wilcoxon signed-rank test. The surrogate data test is often used in the field of nonlinear time-series analysis in order to investigate the existence of the nonlinearity hidden in the original time-series data. If the movement of the return rate or the volatility has any nonlinear properties, this means that our modification into nonlinear versions makes sense. This detail is explained in Section 4. Moreover, we can directly

Corresponding author: Tomoya Suzuki, Ph.D., associate professor, research fields: data mining and financial computing. E-mail: tsuzuki@mx.ibaraki.ac.jp.

compare the prediction accuracies given by the original AR-ARCH model and our modified nonlinear versions. For this comparison, we apply the Wilcoxon signed-rank, and this detail is mentioned in Section 5.

2. AR-ARCH Model

We denote the return rate of the i-th stock at the time t as

$$r_i(t) = \frac{x_i(t) - x_i(t-1)}{x_i(t-1)} \quad (1)$$

First, the AR model is defined as follows:

$$\hat{r}_i(t+1) = \beta_0 + \sum_{k=1}^{p} \beta_k + r_i(t-k+1) \quad (2)$$

where, the AR coefficient $\{\beta_k\}_{k=0}^{p}$ can be decided by the least squares method and the AR order p can be decided by the CV (cross-validation) method or the AIC (Akaike information criterion) to avoid the overfitting problem. Then, the residual $\epsilon_i(t)$ corresponds to the prediction error, and therefore

$$\epsilon_i(t) = r_i(t) - \hat{r}_i(t) \quad (3)$$

If the return rate $r_i(t)$ is the stationary stochastic process, $\epsilon_i(t) \sim N(0, \sigma_i^2)$. However, in financial markets, the volatility $\sigma_i(t)$ temporally changes, and namely $r_i(t)$ is well-known as the non-stationary process.

To model the movement of the volatility $\sigma_i(t)$, the ARCH mode[1] is defined as follows:

$$\hat{\sigma}_i^2(t+1) = \theta_0 + \sum_{k=1}^{q} \theta_i \epsilon_i^2(t-k+1) \quad (4)$$

$$\epsilon_i(t) = \sigma_i(t) u_i(t) \quad (5)$$

Where, $u_i(t) \sim N(0, 1^2)$ and $u_i(t)$ is the residual of the ARCH model. Then, the ARCH coefficient $\{\theta_k\}_{k=0}^{q}$ can be decided by the maximum likelihood estimation, and the ARCH order q can be decided by the CV method or the AIC.

The AR-ARCH model [2, 3] simultaneously applies these two models to predict the future return rate $\hat{r}_i(t+1)$ by the AR model and to predict the future volatility $\hat{\sigma}_i(t+1)$ by the ARCH model under the assumption of $r_i(t+1) \sim N(\hat{r}_i(t+1), \hat{\sigma}_i(t+1))$.

In our study, we use the correlation coefficient C to evaluate the prediction accuracy:

$$C_r = \frac{Cov[r_i(t+1), \hat{r}_i(t+1)]}{\sqrt{Var[r_i(t+1)]}\sqrt{Var[\hat{r}_i(t+1)]}} \quad (6)$$

$$C_\sigma = \frac{Cov[\sigma_i(t+1), \hat{\sigma}_i(t+1)]}{\sqrt{Var[\sigma_i(t+1)]}\sqrt{Var[\hat{\sigma}_i(t+1)]}} \quad (7)$$

However, we cannot observe the true value of $\sigma_i(t+1)$, but can observe the prediction error $\epsilon_i(t+1) = r_i(t+1) - \hat{r}_i(t+1)$. Then, from Eq. (5),

$$\sigma_i^2(t+1) u_i^2(t+1) = \epsilon_i^2(t+1) \quad (8)$$

Here, $\sigma_i^2(t+1)$ is independent of $u_i^2(t+1)$, and therefore

$$E[\sigma_i^2(t+1)] \cdot E[u_i^2(t+1)] = E[\epsilon_i^2(t+1)] \quad (9)$$

Because $E[u_i^2(t+1)] = 1$ and $\epsilon_i^2(t+1)$ has already been observed and fixed,

$$E[\sigma_i^2(t+1)] = \epsilon_i^2(t+1) \quad (10)$$

$$E[\sigma_i(t+1)] = \sqrt{\epsilon_i^2(t+1)} = |\epsilon_i(t+1)| \quad (11)$$

Therefore, we substitute $|\epsilon_i(t+1)|$ for the unobservable $\sigma_i(t+1)$, and modify Eq. (7) into

$$C_\sigma = \frac{Cov[|\epsilon_i(t+1)|, \hat{\sigma}_i(t+1)]}{\sqrt{Var[|\epsilon_i(t+1)|]}\sqrt{Var[\hat{\sigma}_i(t+1)]}} \quad (12)$$

3. Nonlinear Version of AR-ARCH Model

As our scheme, we modify the AR-ARCH model into nonlinear versions by applying a nonlinear prediction method.

First, Eqs. (2) and (4) are generalized as follows:

$$\hat{r}_i(t+1) = \boldsymbol{F}_1[r_i(t), r_i(t-1), \cdots, r_i(t-p+1)] \quad (13)$$

$$\hat{\sigma}_i^2(t+1) = \boldsymbol{F}_2[\epsilon_i^2(t), \epsilon_i^2(t-1), \cdots, \epsilon_i^2(t-p+1)] \quad (14)$$

As shown in Eqs. (2) and (4), the AR and ARCH models are based on linear auto-regressive processes, and therefore $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are linear functions. However, if the movements of the return rate $r(t)$ and its volatility $\sigma(t)$ include nonlinearity, nonlinear time-series models must be better. For this purpose, the present study applies the local-linear approximation method [6], which is one of nonlinear predictions based on the chaotic dynamical theory.

3.1 NAR-ARCH Model

The first modification only focuses the AR part of Eq. (2). Let us denote the input into $\boldsymbol{F}_1$ of Eq. (13) as $\boldsymbol{v}_i(t)$:

$$\boldsymbol{v}_i(t) = [r_i(t), r_i(t-\tau), \cdots, r_i(t-(p-1)\tau)] \quad (15)$$

From the viewpoint of the Takens' embedding theorem, $\boldsymbol{v}_i(t)$ corresponds to the reconstructed attractor. Then, τ is the delay time and p is the embedded dimension.

From the analogy with Eq. (2), we set $\tau = 1$ and optimize p by the CV method.

Next, to predict the future of $\boldsymbol{v}_i(t)$, we estimate $\widehat{\boldsymbol{v}}_i(t+1)$ by learning the function $\boldsymbol{F}_1$ locally. For this purpose, the nearer neighbors from $\boldsymbol{v}_i(t)$ have more important information, and so we calculate the weight parameter as follows:

$$w_i(t,\alpha) = \exp(-bl_i(t,\alpha)) \quad (16)$$

$$l_i(t,\alpha) = |\boldsymbol{v}_i(t) - \boldsymbol{v}_i(t-\alpha)| \quad (17)$$

Then, the weight matrix $\boldsymbol{W}(t)$ is defined by

$$\boldsymbol{W}(t) = \begin{bmatrix} w_i(t,1) & 0 & \cdots & 0 \\ 0 & w_i(t,2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & w_i(t,L-p) \end{bmatrix} \quad (18)$$

Where, L is the length of learning data. Moreover, we denote

$$\boldsymbol{X}(t-1) = \begin{bmatrix} \boldsymbol{v}_i(t-1) & 1 \\ \boldsymbol{v}_i(t-2) & 1 \\ \vdots & \vdots \\ \boldsymbol{v}_i(t-L-p) & 1 \end{bmatrix} \quad (19)$$

$$\boldsymbol{Y}(t) = [r_i(t), r_i(t-1), \cdots, r_i(t-L-p+1)]^{\mathrm{T}} \quad (20)$$

and denote the local structure of the function $\boldsymbol{F}_1$ as $\boldsymbol{F}_1(t)$:

$$\boldsymbol{F}_1(t) = [\beta_0(t), \beta_1(t), \cdots, \beta_p(t)] \quad (21)$$

Then, we can estimate it by minimizing the approximation error of

$\boldsymbol{W}(t)\boldsymbol{Y}(t) = \boldsymbol{W}(t)\boldsymbol{X}(t-1)\boldsymbol{F}_1(t)$, and therefore

$$\widehat{\boldsymbol{F}}_1(t) = [\boldsymbol{X}(t-1)^{\mathrm{T}}\boldsymbol{W}(t)^{\mathrm{T}}\boldsymbol{W}(t)\boldsymbol{X}(t-1)]^{-1} \cdot \boldsymbol{X}(t-1)^{\mathrm{T}}\boldsymbol{W}(t)^{\mathrm{T}}\boldsymbol{W}(t)\boldsymbol{Y}(t) \quad (22)$$

This weighted least square method is a nonlinear approximation because the function $\boldsymbol{F}_1(t)$ can be time-varying and can be approximated by a hypersurface. However, if $b = 0$ in Eq. (16), the weighted parameters are meaningless and then $\boldsymbol{F}_1$ is globally approximated by a hyper plane. Namely, this is the same as the linear approximation of the AR model. For this reason, we set $b = 1$. Finally, by using

$$\widehat{\boldsymbol{Y}}(t+1) = \boldsymbol{X}(t)\widehat{\boldsymbol{F}}_1(t) \quad (23)$$

we can obtain the predicted return rate $\hat{r}_i(t+1)$ from the first element of $\widehat{\boldsymbol{Y}}(t+1)$. We call it the NAR model, where N stands for "Nonlinear."

Next, to predict the volatility $\sigma(t)$, the NAR-ARCH model uses the original ARCH part of Eq. (4) to fit the residual of the NAR model $\epsilon_i(= r_i - \hat{r}_i)$, i.e., the prediction error of the NAR model.

3.2 AR-NARCH Model

The second modification only focuses the ARCH part of Eq. (4). At the beginning, the return rate $r_i(t)$ is predicted by the AR model of Eq. (2), and some prediction errors $\{\epsilon_i(t)\}$ are observed historically as the residuals of the AR model. Next, we fit these prediction errors by the nonlinear approximation used in Section 3.1.

Similarly to Section 3.1, let us denote the input into $\boldsymbol{F}_2(t)$ of Eq. (14) as $\boldsymbol{v}_i(t)$:

$$\boldsymbol{v}_i(t) = [\epsilon_i^2(t), \epsilon_i^2(t-\tau), \cdots, \epsilon_i^2(t-(q-1)\tau)] \quad (24)$$

From the analogy with Eq. (4), we set the delay time τ to one and optimize the embedding dimension q by the CV method.

Next, we apply Eq. (24) to Eqs. (16)-(23), but we modify Eq. (20) into

$$\boldsymbol{Y}(t) = [\sigma_i(t), \sigma_i(t-1), \cdots, \sigma_i(t-L-q+1)]^{\mathrm{T}} \quad (25)$$

$$= [|\epsilon_i(t)|, |\epsilon_i(t-1)|, \cdots, |\epsilon_i(t-L-q+1)|]^{\mathrm{T}} \quad (26)$$

because the true values of $\{\sigma_i(t)\}$ is unknown and so Eq. (11) is applied into $\boldsymbol{Y}(t)$. Moreover, we modify Eqs. (21)-(23) into

$$\boldsymbol{F}_2(t) = [\theta_0(t), \theta_1(t), \cdots, \theta_q(t)] \quad (27)$$

$$\widehat{\boldsymbol{F}}_2(t) = [\boldsymbol{X}(t-1)^{\mathrm{T}}\boldsymbol{W}(t)^{\mathrm{T}}\boldsymbol{W}(t)\boldsymbol{X}(t-1)]^{-1} \cdot \boldsymbol{X}(t-1)^{\mathrm{T}}\boldsymbol{W}(t)^{\mathrm{T}}\boldsymbol{W}(t)\boldsymbol{Y}(t) \quad (28)$$

$$\widehat{\boldsymbol{Y}}(t+1) = \boldsymbol{X}(t)\widehat{\boldsymbol{F}}_2(t) \quad (29)$$

From the first element of $\widehat{\boldsymbol{Y}}(t+1)$, we can obtain the predicted volatility $\hat{\sigma}_i(t+1)$. We call it the NARCH model, where N stands for "Nonlinear." However, in Eq. (16), the NARCH model sets

$b = \frac{1}{\langle|\epsilon_i|\rangle}$ to approximately equalize the scale of weight $w_i(t, \alpha)$ with the case of using Eq. (15).

3.3 NAR-NARCH Model

Finally, we apply the NAR model defined by Eqs. (15)-(23) to predict the return rate r_i, and then its prediction error ϵ_i is used for the NARCH model defined by Eqs.(24), (16)-(19), and (25)-(29) to predict the volatility σ_i. As the result, we can obtain both predicted values: $\hat{r}_i(t+1)$ and $\hat{\sigma}_i(t+1)$.

4. Surrogate Data Test for Nonlinearity of $\boldsymbol{\tau}$ and $|\boldsymbol{\epsilon}|$

To confirm the advantage of using the NAR and/or NARCH model, we examine whether the temporal movements of $r_i(t)$ and $|\epsilon_i(t)|$ have nonlinear property. For this reason, the FS (Fourier shuffled) surrogate data test [7] can be applied.

Table 1 shows the detail of all the stock data used through this study. However, the surrogate data tests are performed only by the learning data as one of the learning processes.

The FS surrogate data are reproduced by randomly shuffling the phases of the Fourier transform of the original time-series of $\{r_i(t)\}_{t=1}^{L}$ and $\{|\epsilon_i(t)|\}_{t=1}^{L}$. Then, each inverse Fourier transformed series is modified so as to have the same frequency distribution as the original data. This FS surrogate data can preserve only linear properties such as the power spectra and the frequency distribution of the original data, but destroys high-order properties such as nonlinearity. If the FS surrogate data test reveals a significant difference between the original data and its surrogate data, we can conclude that nonlinearity is an essential property to identify the original data. In each surrogate data test, we generate 60 surrogate data sets, and apply the nonlinear prediction mentioned in Section 4 to all of the surrogate data sets and the original data. Here, let us denote their prediction accuracies as $C_r(s)$ and $C_\sigma(s)$, where $s = 0$ means the prediction accuracy given by the original data and $s = 1{\sim}60$ means that given by the s-th surrogate data set. Finally, we calculate the z-score by

$$z = \frac{C_\tau - E[C_\tau(s)]}{\sqrt{Var[C_\tau(s)]}} \text{or } z = \frac{C_\sigma - E[C_\sigma(s)]}{\sqrt{Var[C_\sigma(s)]}} \qquad (30)$$

If $z > 1.645$, we can conclude that the original time-series data has nonlinear property by the 5[%] significance level because this nonlinear property is available for the nonlinear prediction, and consequently the prediction accuracy of the original data is better than those of the surrogate data sets.

Tables 2 and 3 show the results. According to Table 2, the return rate r_i hardly has nonlinearity in most markets and periods. This means that the advantage of using the NAR model is not large. However, in American market during the period 2, we can see the rejection of 10 percent. On the other hand, the residual $|\epsilon_i|$ shows more number of rejections especially during the period 1. This means that the residual $|\epsilon_i|$ has some nonlinearity and the NARCH model is useful to predict the volatility σ_i. However, the number of rejections is reduced during the period 2.

5. Predictive Performance of Each Model

Next, by using the test data shown in Table 1 as the out-of-sample data, we investigate the prediction accuracies (C_r and C_σ) of each model. Moreover, we performed the Wilcoxon signed-rank test to examine

Table 1 Details of the stock data used for all simulations.

		Japanese market	American market
	The number of stocks	201 stocks	273 stocks
Period 1	Leaning data	1995/2 - 2000/3 (about 5 years)	1992/1 - 1996/12 (about 5 years)
	Test data	2000/4 - 2005/4 (about 5 years)	1996/1 - 2001/11 (about 5 years)
	The number of stocks	201 stocks	273 stocks
Period 2	Leaning data	1995/2 - 2000/3 (about 5 years)	1992/1 - 1996/12 (about 5 years)
	Test data	2000/4 - 2005/4 (about 5 years)	1996/1 - 2001/11 (about 5 years)

Table 2 The number of stocks rejected by the FS surrogate data test of the return rate r_i.

	Japanese market	American market
Period 1	3 (1.5%)	27 (10%)
Period 2	10 (1.7%)	14 (5%)

Table 3 The same as Table 2, but the FS test of the residual $|\epsilon_i|$ given by the AR or NAR model.

Given by the AR model		
	Japanese market	American market
Period 1	66 (32%)	82 (30%)
Period 2	38 (6%)	7 (2%)

Given by the NAR model		
	Japanese market	American market
Period 1	66 (32%)	81 (30%)
Period 2	46 (8%)	6 (2%)

the improvement by modifying the original AR-ARCH model into our nonlinear versions: the NAR, the AR-NARCH, the NAR-ARCH, and the NAR-NARCH models. If the p-value is less than 0.05, we can conclude that our modification can improve the prediction accuracy.

Table 4 shows the results of the return rate r_i, and Table 5 shows those of the residual $|\epsilon_i|$ derived from the AR or NAR model. Compared to the FS surrogate data tests, rather positive results are shown. In Table 4, we can see the improvement of C_r by the NAR model during the period 1. Moreover, in Table 5, we can see that of C_σ by NARCH model in every market and period. Although the surrogate data tests cannot aggressively support the nonlinearity hidden in real stock data, the prediction accuracy shows the improvement given by our modification into nonlinear versions. This reason is that our nonlinear models use the hyper surface to approximate $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ of Eqs. (13) and (14), and this hypersurface intensively includes the hyperplane approximated by the linear AR and/or ARCH model. In addition to this, Tables 4 and 5 finally conclude that our NAR-NARCH model is the best.

6. Conclusions

We modified the original AR-ARCH model into some nonlinear versions by applying a nonlinear approximation method. To confirm the validity of our modification, we performed the FS surrogate data tests and compared the fitting performance of each model to the learning data. Although the surrogate data tests do not aggressively support the nonlinearity of real stock data, we could confirm that the predictive

Table 4 Prediction accuracy C_r of Eq. (6). Each accuracy is the average value given by all stocks. Then, p shows the p-value evaluated by the Wilcoxon signed-rank test to examine the improvement by modifying the AR model into the NAR model. Each bold figure means that the improvement can be confirmed by the 5[%] significance level (i.e., $p < 0.05$).

	Japanese market		American market	
	AR	NAR	AR	NAR
Period 1	0.0517	0.0528(p = 0.017)	0.0115	0.0159 (p = 0.034)
Period 2	-0.0181	-0.0138(p = 0.189)	0.0069	0.0064 (p = 0.217)

Table 5 The same as Table. 4, but the prediction accuracy C_σ of Eq. (12). Each bold figure ($p < 0.05$) means the improvement by modifying the original AR-ARCH model into our nonlinear versions.

	Japanese market			
	AR-ARCH	AR-NARCH	NAR-ARCH	NAR-NARCH
Period 1	0.1935	0.2059 (p = 0.003)	0.1938 (p = 0.420)	0.2038 (p = 0.000)
Period 2	0.2008	0.2133 (p = 0.000)	0.1718 (p = 1.000)	0.2181 (p = 0.000)

	American market			
	AR-ARCH	AR-NARCH	NAR-ARCH	NAR-NARCH
Period 1	0.1025	0.1388 (p = 0.000)	0.0997 (p = 0.356)	0.1404 (p = 0.000)
Period 2	0.0902	0.1496 (p = 0.000)	0.1175 (p = 0.000)	0.1847 (p = 0.000)

performance of the test data is improved by our modification into nonlinear versions. This improvement was also confirmed by the statistical significant tests, and especially the NAR-NARCH model could realize the most significant improvement.

This research was partially supported by the Grant-in-Aid for Scientific Research (C) (No.25330280) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Engle, R. F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50: 987-1007.

[2] Weiss, A. A. 1986. "Asymptotic Theory for ARCH Models: Estimation and Testing." *Econometric Theory* 2:107-31.

[3] Pantula, S. G. 1988. "Estimation of Autoregressive Models with ARCH Errors." *The Indian Journal of Statistics* 50: 119-38.

[4] Morimoto, K., Saito, M., Inose, S., Kannari, A., and Suzuki, T. 2014. "Principal Component Analysis for the Nonlinear Portfolio Model." *Journal of Signal Processing* 18: 177-80.

[5] Inose, S., Suzuki, T., and Yamanaka, K. 2013. "Nonlinear Portfolio Model and its Rebalance Strategy." NOLTA IEICE4: 351-64.

[6] Farmer, J. D., and Sidorowich, J. J. 1987. "Prediction Chaotic Time Series." *Physical Review Letters* 59: 845-48.

[7] Chang, T., Sauer, T., and Schiff, S. J. 1995. "Tests for Nonlinearity in Short Stationary Time Series." *Chaos* 5: 118-26.