

Estimation of the Vector Autoregressive Model with the Multivariate Skew Normal Distribution for the Shocks: Application to Two Real-World Datasets

Manijeh Mahmoodi 

Ph.D. Candidate in Statistics,
Allameh Tabataba'i University,
Tehran, Iran

Mohammad Reza Salehi Rad* 

Associate Professor, Department
of Statistics, Allameh Tabataba'i
University, Tehran, Iran

Abstract

Modeling plays a crucial role in economic and financial research, forming the foundation for analysis, decision-making, policy development, and planning. Assumptions made during the modeling process are particularly important for estimation and forecasting, as they can significantly influence the results. One of the most widely used classical time series models is the autoregressive model, in which current values are expressed as a finite linear combination of past values. However, in real-world scenarios, many variables interact with each other. To capture these interdependencies, vector time series models—an important class of multivariate time series models—are employed. The vector autoregressive (VAR) models are commonly used in economic and financial modeling. VAR models are typically formulated assuming that the shocks (or noise terms) follow a normal distribution. However, in economic and financial contexts—particularly in macroeconomics—shocks do not often follow a symmetric distribution. The present article focused on a VAR model in which the shocks follow a multivariate skew normal (MSN) distribution. The expectation conditional maximization (ECM) algorithm were used to estimate the model parameters. Finally, using real-world datasets from Canada and Iran—where the shocks exhibit skewness—the study found

* Corresponding Author: E-mail: salehirad@atu.ac.ir

How to Cite: Mahmoodi, M. & Salehi Rad, M.R. (2025). Estimation of the Vector Autoregressive Model with the Multivariate Skew Normal Distribution for the Shocks: Application to Two Real-World Datasets. *Iranian Journal of Economic Research*, 30(102), 38-63.

that the VAR model with MSN-distributed shocks is more efficient than the VAR model with multivariate normal distribution for shocks.

1. Introduction

The multivariate normal distribution is commonly used to model shocks in VAR models. However, in fields such as economics, finance, the stock market, and medicine, various factors can introduce skewness (asymmetry) into the shocks, resulting in non-symmetric distributions. In such cases, the normal distribution becomes an inappropriate choice. To address this, the multivariate skew distribution-which accounts for asymmetry-should be used for modeling shocks. Despite its relevance, this approach has received limited attention in previous research. The family of multivariate skew distributions is broad and complex, posing practical challenges. The present study aimed to test the VAR model in which the shocks follow a multivariate skew normal (MSN) distribution, using the real-world datasets from Canada and Iran.

2. Materials and Methods

Consider the VAR model of order p :

$$\text{VAR}(p): y_t = v + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + \varepsilon_t$$

$$v = \begin{bmatrix} v_{1t} \\ v_{2t} \\ \vdots \\ v_{kt} \end{bmatrix}, \quad y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{kt} \end{bmatrix}, \quad A_i = \begin{bmatrix} a_{11,i} & \dots & a_{1k,i} \\ \vdots & \dots & \vdots \\ a_{k1,i} & \dots & a_{kk,i} \end{bmatrix},$$

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{kt} \end{bmatrix} \quad i = 1, \dots, p$$

where $\varepsilon_t \stackrel{iid}{\sim} MSN(o, \Sigma, S)$, o is location parameter, Σ is the scale parameter, and S is the skew parameter. Its density function is given by:

$$f(\varepsilon_t) = 2^k \phi_k(o, \Omega) \Phi(S' \Omega^{-1} \varepsilon_t; 0, \Delta)$$

where $\Omega = \Sigma + SS'$, $\Delta = (I + S' \Sigma^{-1} S)^{-1} = I - S' \Omega^{-1} S$, $S = \text{diag}(s_1, s_2, \dots, s_k)$, and $\phi_k(\cdot, \cdot)$ denotes density function of MN, while $\Phi(\cdot, \cdot)$ denotes the standard cumulative distribution function.

To find the maximum likelihood estimates of the parameters requires derivatives of the log-likelihood function; however, these derivatives do not have closed-form expressions. Therefore, they must be approximated using numerical methods. The maximum likelihood estimators were obtained via the expectation conditional maximization (ECM) algorithm. Based on the hierarchical representation of the multivariate skew normal distribution, we have:

$$\varepsilon_t | H_t = h_t \sim MN_k(S h_t, \Sigma)$$

$$H_t \sim TN_k(o, I) I(\mathfrak{R}_+^k)$$

$$H_t | \varepsilon_t \sim TN_k(S' \Omega^{-1} \varepsilon_t; \Delta, \mathfrak{R}_+^k)$$

The logarithm of the conditional likelihood function in VAR(p) can be formulated as:

$$\ln L(\Theta | \varepsilon, H) = \ell(\Theta | \varepsilon, H)$$

$$\propto -\frac{T}{2} \ln |\Sigma| - \frac{1}{2} \sum_{t=1}^T \left(Y_t - \sum_{j=1}^p A_j Y_{t-j} - S h_t \right)' \Sigma^{-1} \left(Y_t - \sum_{j=1}^p A_j Y_{t-j} - S h_t \right) - \frac{1}{2} \sum_{t=1}^T h_t' h_t$$

where $\Theta = (A_1, \dots, A_p, \Sigma, S)$. The expectation of the logarithm of the conditional likelihood is denoted by $Q(\Theta | \Theta^{(k)}) = E(\ell(\Theta | y)) \equiv E(\ell(\Theta | \varepsilon))$, and the steps for the maximizing $Q(\Theta | \Theta^{(k)})$ are as below:

- Step 1: Assuming no skewness, estimate the initial values for the coefficients and scale parameters.
- Step 2

$$\hat{A}^{(k+1)}_i = \left[\sum_{t=1}^p \left(Y_t - \sum_{j=1}^p \hat{A}_j^{(k)} Y_{t-j} - \hat{S}^{(k)} E(\widehat{H_t} | \varepsilon_t)^{(k)} \right) Y'_{t-1} \right] \times \left[\sum_{t=1}^T Y_{t-i} Y'_{t-i} \right]^{-1}$$

- Step 3

$$\hat{S}^{(k+1)} = \text{diag} \left(\sum_{t=1}^T \left(Y_t - \sum_{j=1}^p \hat{A}_j^{(k)} Y_{t-j} \right) (\hat{E}(H_t | \varepsilon_t)^{(k)})' \right) \left[\sum_{t=1}^T \hat{E}(H_t H_t' | \varepsilon_t)^{(k)} \right]^{-1}$$

- Step 4

$$\hat{\Sigma}^{(k+1)} = \frac{1}{T} \left[\sum_{t=1}^T \left(Y_t - \sum_{j=1}^p \hat{A}_j^{(k)} Y_{t-j} \right) \left(Y_t - \sum_{j=1}^p \hat{A}_j^{(k)} Y_{t-j} \right)' - \sum_{t=1}^T \left(Y_t - \sum_{j=1}^p \hat{A}_j^{(k)} Y_{t-j} \right) \hat{E}'(H_t | \varepsilon_t)^{(k)} \hat{S}'^{(k)} - \right.$$

$$\left(\sum_{t=1}^T \left(Y_t - \sum_{j=1}^p \hat{A}_j^{(k)} Y_{t-j} \right) \hat{E}'(H_t | \varepsilon_t)^{(k)} \hat{S}'^{(k)} \right)' + \sum_{t=1}^T \hat{S}^{(k)} \hat{E}(H_t H_t' | \varepsilon_t)^{(k)} \hat{S}^{(k)} \right]$$

- Step 5: Repeat steps 2 to 5 until the convergence condition of the algorithm is established:

$$\left| \frac{\ell(\hat{\theta}^{(k+1)} | y)}{\ell(\hat{\theta}^{(k)} | y)} - 1 \right| \leq \varepsilon$$

3. Results and Discussion

The performance of the proposed method was evaluated using two real-world datasets from Canada and Iran. The Dickey-Fuller test was employed to determine the stationarity of the data, while the Akaike Information Criterion (AIC), Hannan-Quinn Criterion (HQC), and Schwartz Bayesian Criterion (BIC) were used to select the order of the VAR model. The Canadian dataset consists of seasonally adjusted employment and unemployment data from 1980 to 2000. According to the Mardia test, the shocks follow a multivariate skew normal (MSN) distribution. Therefore, we estimated the parameters of the VAR(1) model. The AIC and BIC results are presented in Tables 1 and 2, respectively.

Table 1. The Estimated Parameters of Model for Stationary Differenced Canadian Data

Estimation	$\varepsilon_t \sim MN_2$	$\varepsilon_t \sim MSN_2$
$\hat{a}_{11}$	0.9103	0.7094
$\hat{a}_{12}$	0.2139	-0.0149
$\hat{a}_{21}$	-0.2018	-0.4663
$\hat{a}_{22}$	0.3303	0.0255
$\hat{s}_1$	-	0.0703
$\hat{s}_2$	-	0.1066
$\hat{\sigma}_{11}$	0.1646	0.1595
$\hat{\sigma}_{12}$	-0.09106	-0.1003
$\hat{\sigma}_{21}$	0.09106	-0.1003
$\hat{\sigma}_{22}$	0.11400	0.0977

Source: Research findings

Table 2. The AIC and BIC for Data

Distribution	AIC	BIC	-Log-Like
$\varepsilon_t \sim MN_2$	282.0266	286.836	139.011
$\varepsilon_t \sim MSN_2$	260.6781	265.4915	128.339

Source: Research findings

The collected data on agriculture, forestry, and fishing (AFF) and employment of women (EW) in Iran span the years 1991 to 2021. According to the Mardia test, the shocks follow a multivariate skew normal (MSN) distribution. The parameter estimates for the VAR(1) model, along with the AIC and BIC values, are presented in Table 3.

Table 3. The Estimated Parameters of Model for Stationary Differenced Iranian Data

Estimation	$\varepsilon_t \sim MN_2$	$\varepsilon_t \sim MSN_2$
$\hat{\delta}_1$	-	0.1104
$\hat{\delta}_2$	-	-0.0798
$\hat{\sigma}_{11}$	1.2397	1.2267
$\hat{\sigma}_{12}$	0.2748	0.2836
$\hat{\sigma}_{21}$	0.2748	0.2836
$\hat{\sigma}_{22}$	0.3855	0.3796
AIC	146.1145	145.9185
BIC	148.8491	148.6531

Source: Research findings

The fitted model can be formulated as follows:

$$y_{1t} = -0.25626 y_{1t-1} - 0.78496 y_{2t-1} + \varepsilon_{1t}$$

$$y_{2t} = -0.09239 y_{1t-1} + 0.29052 y_{2t-1} + \varepsilon_{2t}$$

where y_1 and y_2 represent AFF and EW, respectively.

According to the AIC and BIC criteria presented in Tables 2 and 3 for the Canadian and Iranian data, the VAR model with MSN-distributed shocks is more appropriate than the VAR model with MN-distributed shocks.

4. Conclusion

Considering the VAR(p) model with shocks following a multivariate skew normal (MSN) distribution, the present study employed the maximum likelihood method and the ECM algorithm to estimate the model parameters. Based on two real-world datasets from Canada and Iran, the findings showed that the VAR model with MSN-distributed shocks provides a better fit than the model with MN shocks when the shocks exhibit skewness.

Keywords: Vector Autoregressive, Skewness, Multivariate Skew Normal, Maximum Likelihood Estimation, Expectation Conditional Maximization Algorithm

JEL Classification C13, C32, C51, C46



برآورد مدل اتورگرسو بر داری با توزیع نرمال چوله چندمتغیره برای شوک‌ها و تحلیل عملکرد آن برای دو مجموعه داده واقعی

منیژه محمودی

دانشجوی دکتری رشته آمار، دانشگاه علامه طباطبائی، تهران، ایران

محمدرضا صالحی‌راد *

دانشیار گروه آمار، دانشکده آمار، ریاضی و رایانه، دانشگاه علامه

طباطبائی، تهران، ایران

چکیده

مدل‌سازی مبحث بسیار مهمی در پژوهش‌های اقتصادی و مالی است و نقش بسیار بالایی در تحلیل‌ها و اتخاذ تصمیمات و سیاست‌گذاری و برنامه‌ریزی‌ها دارد. از طرفی در مدل‌سازی‌ها، مفروضات نقش مهمی در مسئله برآوردها و پیش‌بینی‌ها ایفا می‌کنند زیرا می‌توانند بر نتایج مدل‌ها و تحلیل‌ها اثرگذار باشند. یکی از پرکاربردترین مدل‌های سری زمانی کلاسیک، مدل اتورگرسو است که مقادیر فعلی فرآیند ترکیب خطی متناهی از مقادیر گذشته آن می‌باشد. از طرف دیگر، در مسائل واقعی متغیرهای زیادی بر هم تأثیر می‌گذارند، به همین دلیل مدل‌های سری زمانی برداری به کار گرفته می‌شوند که جزء سری زمانی چندمتغیره محسوب می‌گردند. مدل اتورگرسو برداری در مدل‌سازی‌های اقتصادی و مالی بسیار پرکاربرد است. از سوی دیگر، مدل‌های اتورگرسو برداری معمولاً با شوک‌های (نویز) نرمال در نظر گرفته می‌شوند. از آن‌جا که در مسائل اقتصادی و مالی به‌ویژه اقتصاد کلان، شوک‌ها حالت تقارن ندارند، در این مقاله مدل اتورگرسو برداری با توزیع نرمال چوله چندمتغیره روی شوک‌ها در نظر گرفته می‌شود و چون برآورد پارامترها مرحله مهمی در مدل‌سازی است، برآورد پارامترهای مدل با استفاده از الگوریتم بیشینه‌سازی امید ریاضی شرطی به‌دست آورده می‌شوند. در پایان با استفاده از دو مجموعه داده‌های واقعی کانادا و ایران که شوک‌ها دارای چولگی هستند و براساس معیارهای ارزیابی مدل‌ها، نشان داده می‌شود که مدل اتورگرسو برداری با توزیع نرمال چوله چندمتغیره برای شوک‌ها در این داده‌ها کارایی بیشتری نسبت به مدل اتورگرسو با توزیع نرمال چندمتغیره دارد.

کلیدواژه‌ها: اتورگرسو برداری، الگوریتم بیشینه‌سازی امید ریاضی شرطی، برآورد بیشینه درست‌نمایی، چولگی، نرمال چوله چندمتغیره

طبقه‌بندی JEL: C13، C32، C51، C46

* نویسنده مسئول : salelehirad@atu.ac.ir

۱. مقدمه

مدل‌سازی یکی از ابزارهای مهم در مسائل اقتصادی و مالی جهت تصمیم‌گیری‌ها و سیاست‌گذاری‌های مدیریتی است. از پرکاربردترین روش‌های مدل‌سازی در زمینه‌های اقتصادی و مالی می‌توان به سری‌های زمانی اشاره کرد. سری‌های زمانی یکی از زیرمجموعه‌های فرآیندهای تصادفی است. باکس و جنکینز^۱ در سال ۱۹۷۰ مدل‌های سری زمانی کلاسیک را معرفی کردند. یکی از معروف‌ترین مدل‌های سری زمانی کلاسیک، مدل اتورگرسیو^۲ (AR) است که در آن تنها یک متغیر وجود دارد و اثر گذشته متغیر روی آن بررسی می‌شود و با فرض توزیع نرمال برای شوک‌ها پارامترها برآورد می‌شوند. از طرف دیگر، به دلیل تأثیر عوامل (متغیرهای کمکی) متعدد بر پدیده مورد بررسی، از مدل‌های کلاسیک استفاده نمی‌شود و سری‌های زمانی چندمتغیره به کار گرفته می‌شوند. تسای^۳ (۲۰۰۲) به معرفی و تحلیل مدل سری زمانی در حالت یک‌متغیره و چندمتغیره در زمینه‌های مالی و اقتصادی و لوتکیل^۴ (۲۰۰۵) به بررسی مدل‌های اتورگرسیو برداری، میانگین متحرک برداری و ترکیب آن‌ها پرداختند. مدل اتورگرسیو برداری ساختاری^۵ (SVAR) توسط کیلیان و لوتکیل^۶ (۲۰۱۷) و لوتکیل (۲۰۲۰) مورد تحلیل و بررسی قرار گرفت. مدل‌های سری زمانی چندمتغیره در بسیاری از زمینه‌های اقتصادی، پزشکی، زیست‌شناسی، علوم اجتماعی و... استفاده می‌شود. یکی از مهم‌ترین این مدل‌ها، مدل اتورگرسیو برداری است. اکنون به پیشینه تحقیق در زمینه مدل‌های اتورگرسیو و اتورگرسیو برداری با توزیع شوک‌ها اشاره می‌کنیم.

سیمز^۷ (۱۹۸۰) روش مدل‌سازی بدون ایجاد محدودیت روی ضرایب را به عنوان جایگزینی برای مدل‌سازی در مقیاس بزرگ در مسائل اقتصادسنجی بیان کرد. از نظر او اگر بین متغیرها الگوی همزمانی وجود داشته باشد، باید متغیرها را درون‌زا نگرست و از مدل‌های اتورگرسیو برداری استفاده نمود و در ساخت مدل‌های بزرگ محدودیت ایجاد

-
1. Box, G. & Jenkins, G.
 2. Autoregressive
 3. Tsays, R.S.
 4. Lütkepohl, H.
 5. Structural Vector Autoregressive
 6. Kilian, L. & Lütkepohl, H.
 7. Sims, C.

نکرد. پوراحمدی^۱ (۲۰۰۱) مدل اتورگرسو مرتبه اول را با در نظر گرفتن توزیع‌های نمایی، گاما و هندسی مختلط برای شوک‌ها تحلیل نمود. طارمی و پوراحمدی^۲ (۲۰۰۳) مدل اتورگرسو مرتبه p را با توزیع تی‌استودنت برای نوسانات مورد بررسی قرار دادند. وی^۳ (۲۰۰۶) روش‌های تک‌متغیره و چندمتغیره را در تحلیل سری‌های زمانی مورد بحث قرار داد. مطالعاتی روی مدل اتورگرسو با توزیع نرمال اپسیلون چوله برای نوسانات صورت گرفت (Bondon, 2009). زمانی و سیاره^۴ (۲۰۱۵) استنتاج آماری را در مدل اتورگرسو با خطاهای نامنفی مورد مطالعه قرار دادند. شرفی و نعمت‌اللهی^۵ (۲۰۱۶) مدل اتورگرسو مرتبه اول را با خطاهای چوله بررسی نمودند. لیو و همکاران^۶ (۲۰۱۶) به تحلیل مدل اتورگرسو با توزیع تی پرداختند. قاسمی و همکاران^۷ (۲۰۱۹) فرآیندهای اتورگرسو را با توزیع هندلولی تعمیم‌یافته برای نوسانات مورد بحث قرار دادند. لیو و همکاران^۸ (۲۰۱۹) با استفاده از الگوریتم بیشینه‌سازی امید ریاضی پارامترهای مدل را با توزیع دم پهن و داده‌های گمشده برآورد کردند. پژوهشگرانی با در نظر گرفتن توزیع نرمال چوله تعمیم‌یافته برای نوسانات، به برآورد مدل اتورگرسو پرداختند (Neethling, et al., 2020).

با گذشت زمان، مدل‌های سری زمانی چندگانه برای مدل‌سازی‌های چندمتغیره معرفی شده و مورد مطالعه قرار گرفتند (Lütkepohl, 1991 & 2005). پژوهشگرانی مدل‌های اتورگرسو برداری را بررسی کرده و تفاوت بین اتورگرسو کاهش‌یافته، اتورگرسو بازگشتی و اتورگرسو ساختاری را مورد بحث قرار دادند (Stock & Watson, 2001). شیوه‌های سری زمانی چندمتغیره بیزی برای مدل‌های اتورگرسو برداری در اقتصاد کلان تجربی توسط محققان مورد مطالعه قرار گرفته است (Koop & Korobilis, 2009). پژوهش‌هایی جهت برآورد بیزی برای مدل‌های اتورگرسو برداری و برآورد آن مدل‌ها و همچنین پیش‌بینی مدل با نوسانات تصادفی با توزیع دم پهن نیز صورت گرفته است (Ni

-
1. Pourahmadi, M.
 2. Tarami, B. & Pourahmadi, M.
 3. Wei, W.
 4. Zamani Mehryan, S. & Sayyareh, A.
 5. Sharafi, M. & Nematollahi, A.
 6. Liu, Y., et al.
 7. Ghasemi, S., et al.
 8. Liu, Y., et al.

(Wai, et al., 2017; & Sun, 2015). ملکی و همکاران^۱ (۲۰۱۹) یک توزیع دم پهن نامتقارن را برای مدل اتورگرسیو برداری در نظر گرفته و مورد بررسی قرار دادند. مدل اتورگرسیو برداری امکان بررسی همزمان چند متغیر را می‌دهد بنابراین در پژوهش‌های اخیر در زمینه‌های اقتصادی و مالی بسیار مورد توجه قرار گرفته است. حیدری^۲ (۲۰۱۱) مدل اتورگرسیو برداری و بیزی را برای پیش‌بینی تورم به کار برده است. محمدی و همکاران^۳ (۱۳۹۸) پژوهش پویایی‌های کلان اقتصادی مقررات‌زدایی در بازارهای محصول کار در کشورهای منافع^۴ را با مدل اتورگرسیو پنل انجام دادند. خرسندی^۵ و همکاران (۲۰۲۲) آثار شوک‌های اقتصادی خارجی بر متغیرهای کلان اقتصادی ایران را با رویکرد اتورگرسیو برداری جهانی^۶ (GVAR) مورد مطالعه قراردادند. مدل GVAR مدل‌های تصحیح خطای برداری داخلی را که در آن متغیرهای داخلی با متغیرهای خارجی خاص هر کشور مرتبط هستند، ترکیب می‌کند. متغیرهای خارجی خاص هر کشور از متغیرهای داخلی ساخته شده است تا بتواند با تجارت بین‌الملل یا الگوی مورد نظر کشور مورد بررسی مطابقت داشته باشد. تحلیل شبکه بازار مسکن بین استان‌های ایران توسط میرزایی و همکاران^۷ (۲۰۲۳) با استفاده از مدل اتورگرسیو برداری مورد تجزیه و تحلیل قرار گرفت.

با توجه به مطالعات انجام شده بر روی مدل‌های اتورگرسیو برداری، شوک‌های مدل اغلب دارای توزیع نرمال چندمتغیره هستند (Kilian & Lütkepohl, 2017). توزیع نرمال چندمتغیره یکی از پرکاربردترین توزیع‌های متقارن است. در برخی موارد در مسائل واقعی در زمینه‌های اقتصادی، مالی، بورس، پزشکی، زیست‌شناسی و... تقارن روی شوک‌ها وجود ندارد. به همین دلیل باید از توزیع نامتقارن (چوله) برای شوک‌ها استفاده نمود. از طرف دیگر، خانواده توزیع‌های چوله چندمتغیره بسیار بزرگ است و کار روی آن‌ها آسان نیست لذا در این مقاله توزیع نرمال چوله چندمتغیره در نظر گرفته شده و پارامترهای مدل اتورگرسیو برداری با استفاده از الگوریتم بیشینه‌سازی امید ریاضی شرطی برآورد شده است.

1. Maleki, M., et al.

2. Heidari, H.

3. Mohammadi, T., et al.

4. Middle East and North Africa (MENA) منطقه خاورمیانه و شمال آفریقا

5. Khorsandi, M., et al.

6. Global Vector Autoregressive (GVAR)

7. Mirzaei, H., et al.

سپس با استفاده از معیارهای ارزیابی مدل روی داده‌های واقعی با شوک‌های نامتقارن نشان داده می‌شود که مدل اتورگرسیو برداری با توزیع نرمال چوله چندمتغیره برای شوک‌ها در این داده‌ها نسبت به مدل با توزیع نرمال چندمتغیره برای شوک‌ها مناسب‌تر است. مقاله به شرح زیر تنظیم شده است:

در بخش ۲، پیشینه پژوهشی توزیع نرمال چوله چندمتغیره و ویژگی‌های این توزیع معرفی شده است. در بخش ۳، جزئیات مدل اتورگرسیو برداری و الگوریتم پیشینه‌سازی امید ریاضی شرطی برای برآورد پارامترهای مدل شرح داده می‌شود. سپس در بخش ۴ عملکرد شیوه استفاده شده روی داده‌های واقعی ارزیابی و نتیجه‌گیری در بخش ۵ ارائه می‌شود.

۲. مبانی نظری

در این بخش، ابتدا پیشینه پژوهشی توزیع نرمال چوله چندمتغیره و مفاهیم نظری مربوط به آن شرح داده می‌شود.

۱-۲. پیشینه توزیع نرمال چوله چندمتغیره

اولین پیشنهادات برای توزیع‌های نامتقارن به قرن نوزدهم برمی‌گردد و آزالینی^۱ (۱۹۸۵) یکی از پژوهشگرانی است که توزیع نرمال یک‌متغیره را معرفی کرده است. در نظر گرفتن خانواده‌های نامتقارن نرمال توسط محققان زیادی مورد مطالعه قرار گرفته است. معیارهای اندازه‌گیری چولگی و برجستگی توسط ماردیا^۲ (۱۹۷۰) ارائه شده است. آزالینی و کاپیتانیو^۳ (۱۹۹۹) توزیع نرمال چوله چندمتغیره و کاربرد آماری آن را معرفی کرده‌اند و در سال ۲۰۱۴ تحقیقات کامل خود را بر روی نرمال چوله و خانواده آن‌ها ارائه کردند و توزیع نرمال تعمیم‌یافته چوله^۴ (SGN) آن توسط آرلانو واله و همکاران^۵ (۲۰۰۴) معرفی شده است. آزالینی (۲۰۰۵) در مقاله مروری خود به توزیع نرمال چوله یک‌متغیره، چندمتغیره و خانواده توزیع‌های چندمتغیره مرتبط پرداخته است. آرلانو واله و جنتون^۶ (۲۰۰۵) یک کلاسی جدید

1. Azzalini, A.

2. Mardia

3. Azzalini, A. & Capitanio, A.

4. Skew Generalizes Normal

5. Arellano-Valle, R.B., et al.

6. Arellano-Valle, R.B. & Genton, M.G.

از خانواده توزیع نرمال چوله چندمتغیره را به نام توزیع FUSN^۱ معرفی کردند و گسترش آن را به توزیع‌های چوله بیضوی و چوله کروی مورد بحث قرار دادند.

همچنین پژوهشگرانی مانند آرلانوواله و آزالینی^۲ (۲۰۰۶) مطالعاتی در زمینه توزیع نرمال چوله چندمتغیره انجام داده‌اند. ما در این مقاله شوک‌های مدل اتورگرسیو برداری^۳ (VAR) را توزیع نرمال چوله چندمتغیره در نظر می‌گیریم که این مدل در بخش بعد معرفی می‌شود.

۲-۲. توزیع نرمال چوله چندمتغیره

در این بخش قصد داریم به توزیع نرمال چوله چندمتغیره^۴ (MSN) و ویژگی‌های آن بپردازیم. اگر Y یک بردار با اندازه $k \times 1$ از توزیع نرمال چوله چندمتغیره $(Y \sim MSN(\xi, \Sigma, S))$ باشد که ξ بردار مکانی $k \times 1$ ، Σ ماتریس پراکندگی و مثبت مرتبه $k \times k$ و S ماتریس چولگی قطری باشد آن‌گاه طبق مطالعاتی که توسط گوپتا و چانگ^۵ (۲۰۰۲)، سوچیت و همکاران^۶ (۲۰۰۳) و تی سانگ^۷ (۲۰۰۹) در مورد توزیع نرمال چوله چندمتغیره صورت گرفته است، تابع چگالی آن به صورت زیر تعریف می‌شود:

$$f(y) = 2^k \phi_k(0, \Omega) \Phi(S' \Omega^{-1}(y - \xi); 0, \Delta) \quad (1)$$

که در آن $\Delta = (I + S' \Sigma^{-1} S)^{-1} = I - S' \Omega^{-1} S$ ، $\Omega = \Sigma + S S'$

نرمال است. اگر S برابر ماتریس صفر باشد آن‌گاه Y دارای توزیع نرمال چندمتغیره است و با نماد $Y \sim MN(\xi, \Sigma, S)$ نشان داده می‌شود. فرم سلسله‌مراتبی توزیع نرمال چوله چندمتغیره طبق توضیحات زیر در رابطه (۲) نشان داده می‌شود.

سوچیت و همکاران (۲۰۰۳) طبق بررسی توزیع نرمال چوله چندمتغیره در رگرسیون بیزی، آرلانوواله و جنتون^۸ (۲۰۰۵) و تی سانگ و همکاران^۹ (۲۰۰۹) در تحلیل مدل‌ها با

-
1. Fundamental Skew Normal
 2. Arellano-Valle, R.B. & Azzalini, A.
 3. Vector Autoregressive
 4. Multivariate Skew Normal
 5. Gupta, A. & Chang, C.
 6. Sujit, S., et al.
 7. Tsung, L.
 8. Arellano-Valle, R.B. & Genton, M.G.
 9. Tsung, L., et al.

توزیع نرمال چوله، نشان دادند که اگر $Z_0 \sim N_k(O, I_k)$ و $Z_1 \sim N_k(\xi, \Sigma)$ مستقل باشند، آن‌گاه $Y = S|Z_0| + Z_1$ دارای یک توزیع نرمال چوله چندمتغیره خواهد بود و با نماد $Y \sim MSN_k(\xi, \Sigma, S)$ نشان داده می‌شود. همچنین فرم سلسله‌مراتبی توزیع نرمال چوله چندمتغیره برابر است با:

$$Y|W = w \sim MN_k(Sw + \xi, \Sigma) \quad (۲)$$

$$W \sim TN_k(o, I) I(\mathcal{R}_+^k), \quad \mathcal{R}_+^k = (0, +\infty)^k$$

که در آن w یک بردار $k \times 1$ و هر مؤلفه آن دارای توزیع نرمال در بازه صفر و یک (توزیع نرمال کاسته شده) است. طبق توصیف ذکر شده در ساختار فرم سلسله‌مراتبی، اگر $|Z_0| \sim TN(O, I_k)$ آنگاه $E|Z_0| = \sqrt{2/\pi} 1_k$ و ماتریس کوواریانس آن $\text{Var}(|Z_0|) = \left(1 - \frac{2}{\pi}\right) I_k$ خواهد بود و طبق تحقیقات آزالینی و کاپیتانیو^۱ (۲۰۱۴) و آرلانو واله و همکاران^۲ (۲۰۰۵) داریم:

$$E(Y) = E(S|Z_0| + Z_1) = S E|Z_0| + E(Z_1) = S \sqrt{\frac{2}{\pi}} 1_k + \xi \quad (۳)$$

$$\text{Var}(Y) = S \text{Var}|Z_0| S' + \text{Var}(Z_1) = S \left(1 - \frac{2}{\pi}\right) I_k S' + \Sigma = \left(1 - \frac{2}{\pi}\right) S S' + \Sigma \quad (۴)$$

۳. برآورد مدل VAR با توزیع نرمال چوله چندمتغیره برای شوک‌ها

در این بخش، ابتدا مدل اتورگرسیو برداری را بیان کرده و سپس شیوه برآورد پارامترهای مدل با استفاده از الگوریتم بیشینه‌سازی امید ریاضی شرطی شرح داده می‌شود.

۳-۱. مدل اتورگرسیو برداری

در این زیربخش، ما مدل اتورگرسیو برداری مرتبه p با توزیع نرمال چوله چندمتغیره روی شوک‌ها را در نظر می‌گیریم. این مدل به صورت زیر نوشته می‌شود:

$$\text{VAR}(p): y_t = v + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + \varepsilon_t \quad (۵)$$

$$v = \begin{bmatrix} v_{1t} \\ v_{2t} \\ \vdots \\ v_{kt} \end{bmatrix}, \quad y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{kt} \end{bmatrix}, \quad A_i = \begin{bmatrix} a_{11,i} & \dots & a_{1k,i} \\ \vdots & \dots & \vdots \\ a_{k1,i} & \dots & a_{kk,i} \end{bmatrix}, \quad \text{که در آن}$$

1. Azzalini, A. & Capitanio, A.

2. Arellano-Valle, R.B., et al.

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{kt} \end{bmatrix} \quad i = 1, \dots, p$$

و $\varepsilon_t \stackrel{iid}{\sim} MSN(o, \Sigma, S)$ که بردار مکانی، Σ ماتریس پراکندگی (پارامتر مقیاس) با درایه‌های σ_{ij} در سطر i ام و ستون j ام و S ماتریس قطری (پارامتر چولگی)، می‌باشند و تابع چگالی برابر است با:

$$f(\varepsilon_t) = 2^k \phi_k(o, \Omega) \Phi(S' \Omega^{-1} \varepsilon_t; 0, \Delta) \quad (۶)$$

که در آن $\phi_k(\cdot, \cdot)$ تابع چگالی نرمال و $\Phi(\cdot, \cdot)$ تابع توزیع تجمعی نرمال است.

۳-۲. روش بیشینه درست‌نمایی

طبق رابطه ۵، تابع درست‌نمایی با T مشاهده مستقل با فرض بردار مقادیر ثابت v برابر صفر در مدل اتورگرسیون برداری مرتبه p با توزیع نرمال چوله چندمتغیره روی شوک‌ها، به صورت زیر تعریف می‌شود:

$$L(A_1, \dots, A_p, \Omega, S) = 2^{kT} \prod_{t=1}^T \phi_k(y_t - A_1 y_{t-1} - A_2 y_{t-2} - \dots - A_p y_{t-p}; \Omega) \times \Phi(S' \Omega^{-1} (y_t - A_1 y_{t-1} - A_2 y_{t-2} - \dots - A_p y_{t-p})) \quad (۷)$$

و لگاریتم تابع درست‌نمایی برابر است با:

$$\ell(A_1, \dots, A_p, \Omega, S) = kT \ln 2 + \sum \ln \phi_k(y_t - A_1 y_{t-1} - A_2 y_{t-2} - \dots - A_p y_{t-p}; \Omega) + \sum \ln \Phi(S' \Omega^{-1} (y_t - A_1 y_{t-1} - A_2 y_{t-2} - \dots - A_p y_{t-p})) \quad (۸)$$

برای بدست آوردن برآورد بیشینه درست‌نمایی برای پارامترهای ضرایب، چولگی و ماتریس پراکندگی مدل ۵، ما به مشتق‌گیری از رابطه (۸) نیازمندیم. به دلیل این که از حل مشتقات، فرم بسته‌ای برای برآورد پارامترها حاصل نمی‌شود، لازم است از شیوه‌های عددی برای تقریب برآورد استفاده شود. از طرف دیگر، برآورد بیشینه درست‌نمایی براساس شیوه‌های عددی به مقادیر اولیه بسیار حساس است. برای حل این مشکل می‌توان از الگوریتم بیشینه‌سازی امید ریاضی شرطی^۱ (ECM) که در همگرایی نیز پایدارتر است استفاده کرد. براساس فرم سلسله‌مراتبی داریم:

$$\varepsilon_t | H_t = h_t \sim MN_k(S h_t, \Sigma) \quad (۹)$$

$$H_t \sim TN_k(o, I) I(\mathcal{R}_+^k)$$

1. Expectation Conditional Maximization Algorithm (ECM)

$$H_t | \varepsilon_t \sim TN_k (S' \Omega^{-1} \varepsilon_t; \Delta, \mathfrak{R}_+^k)$$

تابع درست‌نمایی شرطی $(L(\theta | \varepsilon, H))$ و لگاریتم تابع درست‌نمایی شرطی به صورت زیر نوشته می‌شود:

$$\begin{aligned} L(\theta | \varepsilon, H) &= \prod_{t=1}^T \phi_k(\varepsilon_t; S h_t, \Sigma) TN(h_t; O, I) \\ &= \prod_{t=1}^T (2\pi)^{-\frac{k}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} (\varepsilon_t - S h_t)' \Sigma^{-1} (\varepsilon_t - S h_t)\right\} 2^k \exp\left\{-\frac{1}{2} h_t' h_t\right\} \end{aligned}$$

و داریم:

$$\begin{aligned} \ln L(\theta | \varepsilon, H) &= \ell(\theta | \varepsilon, H) \\ &\propto \ln \left[\prod_{t=1}^T |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} (\varepsilon_t - S h_t)' \Sigma^{-1} (\varepsilon_t - S h_t)\right\} 2^k \exp\left\{-\frac{1}{2} h_t' h_t\right\} \right] \\ &\propto -\frac{T}{2} \ln |\Sigma| - \frac{1}{2} \sum_{t=1}^T (\varepsilon_t - S h_t)' \Sigma^{-1} (\varepsilon_t - S h_t) - \frac{1}{2} \sum_{t=1}^T h_t' h_t \\ &\propto -\frac{T}{2} \ln |\Sigma| - \frac{1}{2} \sum_{t=1}^T \left(Y_t - \sum_{j=1}^p A_j Y_{t-j} - S h_t \right)' \Sigma^{-1} \left(Y_t - \sum_{j=1}^p A_j Y_{t-j} - S h_t \right) - \frac{1}{2} \sum_{t=1}^T h_t' h_t \end{aligned} \quad (10)$$

که $\theta = (A_1, \dots, A_p, \Sigma, S)$ است.

۳-۳. الگوریتم ECM

در این بخش به معرفی مراحل الگوریتم ECM برای برآورد پارامترهای مدل اتورگرسیو برداری با در نظر گرفتن توزیع نرمال چوله چندمتغیره برای شوک‌ها می‌پردازیم. امید ریاضی لگاریتم درست‌نمایی شرطی با نماد زیر نشان داده می‌شود:

$$Q(\theta | \hat{\theta}) = E(\ell(\theta | y)) \equiv E(\ell(\theta | \varepsilon))$$

و مراحل پیشینه‌سازی شامل گام‌های زیر است:

گام اول: برآورد مقادیر اولیه پارامترهای ضرایب و مقیاس با فرض عدم چولگی، به عبارت دیگر برآورد پارامترهای مدل اتورگرسیو برداری با فرض توزیع نرمال چوله چندمتغیره برای شوک‌ها

گام دوم: برآورد ضرایب مدل

$$\begin{aligned} \hat{A}^{(k+1)}_i &= \left[\sum_{t=1}^p (Y_t - \sum_{j \neq i}^p \hat{A}^{(k)}_j Y_{t-j} - \hat{S}^{(k)} E(\widehat{H_t} | \varepsilon_t)^{(k)}) Y'_{t-1} \right] \times \\ &\quad \left[\sum_{t=1}^T Y_{t-i} Y'_{t-i} \right]^{-1} \end{aligned} \quad (11)$$

که در آن

$$h_t^{(k)} | \varepsilon_t \sim \text{TN}(\widehat{S^{(k)}} \widehat{\Omega^{(k)}} \varepsilon_t, \widehat{\Delta^{(k)}})$$

$$\widehat{\Omega}^{(k)} = \widehat{\Sigma}^{(k)} + (\widehat{S}^{(k)})' (\widehat{S}^{(k)})'$$

$$\Delta^{(k)} = I - (\widehat{S}^{(k)})' (\widehat{\Omega}^{(k)})^{-1} \widehat{S}^{(k)}$$

گام سوم: برآورد پارامتر چولگی

$$\widehat{S}^{(k+1)} = \text{diag} \left(\sum_{t=1}^T (Y_t - \sum_{j=1}^p \widehat{A}_j^{(k)} Y_{t-j}) (\widehat{E}(H_t | \varepsilon_t)^{(k)})' \right) \left[\sum_{t=1}^T \widehat{E}(H_t H_t' | \varepsilon_t)^{(k)} \right]^{-1} \quad (12)$$

گام چهارم: برآورد پارامتر مقیاس

$$\begin{aligned} \widehat{\Sigma}^{(k+1)} = & \frac{1}{T} \left[\sum_{t=1}^T (Y_t - \sum_{j=1}^p \widehat{A}_j^{(k)} Y_{t-j}) (Y_t - \sum_{j=1}^p \widehat{A}_j^{(k)} Y_{t-j})' - \right. \\ & \sum_{t=1}^T (Y_t - \sum_{j=1}^p \widehat{A}_j^{(k)} Y_{t-j}) \widehat{E}'(H_t | \varepsilon_t)^{(k)} \widehat{S}^{(k)} - \\ & \left. \left(\sum_{t=1}^T (Y_t - \sum_{j=1}^p \widehat{A}_j^{(k)} Y_{t-j}) \widehat{E}'(H_t | \varepsilon_t)^{(k)} \widehat{S}^{(k)} \right)' + \sum_{t=1}^T \widehat{S}^{(k)} \widehat{E}(H_t H_t' | \varepsilon_t)^{(k)} \widehat{S}^{(k)} \right] \end{aligned}$$

گام پنجم: تکرار گام‌های دوم الی پنجم تا زمان برقراری شرط همگرایی فرآیند

$$\left| \frac{\ell(\widehat{\theta}^{(k+1)} | y)}{\ell(\widehat{\theta}^{(k)} | y)} - 1 \right| \leq \varepsilon$$

۴. داده‌های کاربردی

در این بخش، عملکرد روش پیشنهادی بر روی دو مجموعه داده‌های واقعی کانادا و ایران با استفاده از نرم‌افزار R بررسی می‌شود. برای تعیین مانایی داده‌ها از آزمون دیکی-فولر^۱ و برای تعیین مرتبه مدل اتورگرسیو برداری ملاک‌های آکائیکه^۲ (AIC)، حنان-کوئین^۳ (HQ) و شوارتز-بیزی^۴ (SC) استفاده می‌شود. لازم به ذکر است برای شرط ایستایی در

1. Dickey-Fuller

2. Akaike

3. Hannan-Quinn

4. Schwarz

مدل اتورگرسیو برداری باید ریشه‌های معادله $|I - Az| = 0$ برای $|z| \geq 1$ برقرار باشد (Lütkepohl, 2005) و این شرط نیز بررسی می‌شود. اگرچه چولگی در حالت یک‌متغیره عدم تقارن است و مفهوم ساده‌ای به نظر می‌رسد اما برای حالت چندمتغیره ماردیا (۱۹۷۴) و بالاکریشنان و اسکارپا^۱ (۲۰۱۲) معیار Mardia را برای بررسی چولگی و کشیدگی نرمال چندمتغیره ارائه کرده‌اند که چولگی چندمتغیره در حالت دوبردار p بعدی به صورت زیر تعریف می‌شود:

$$\gamma_{1,p} = E[(X - \mu)' \Sigma^{-1}(Y - \mu)]^3$$

با بررسی چولگی شوک‌ها، مدل برازش داده می‌شود. مقدار اپسیلون در شرط همگرایی الگوریتم در این بخش 10^{-4} در نظر گرفته شده است. سپس معیار اطلاعات آکائیکه یعنی $BIC = 2k - 2\ell(\hat{\theta} | Y)$ (آکائیک، ۱۹۷۴) و معیار اطلاعات بیزی یعنی $AIC = 2k - 2\ell(\hat{\theta} | Y)$ (شوارتز، ۱۹۷۸) برای ارزیابی عملکرد مدل اتورگرسیو برداری با فرض توزیع نرمال چندمتغیره و توزیع نرمال چوله چندمتغیره برای شوک‌ها بر دو مجموعه داده‌های واقعی کانادا و ایران استفاده می‌شود.

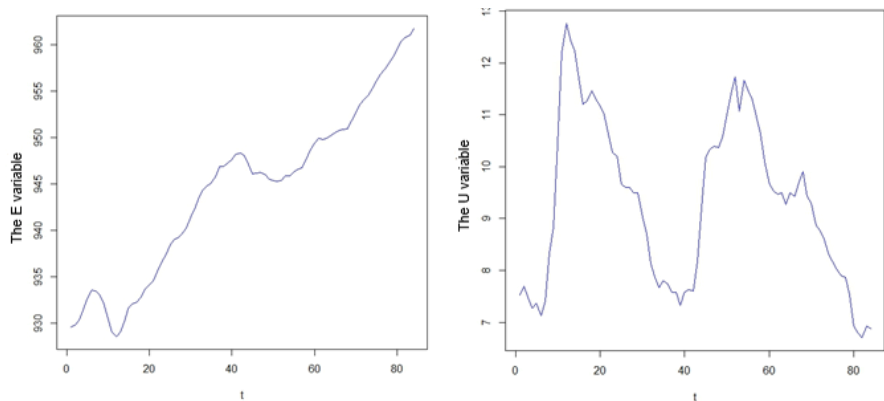
۴-۱. داده‌های کاربردی کانادا

در این بخش، مجموعه‌ای از داده‌های اقتصادی مربوط به کانادا از سال ۱۹۸۰ تا ۲۰۰۰ با متغیر اشتغال (E) و نرخ بیکاری (U) به صورت فصلی در نرم‌افزار R در نظر گرفته شده و به برازش مدل می‌پردازیم. نمودارهای سری زمانی برای داده‌های E و U در مقابل زمان در شکل ۱ رسم شده‌اند. این نمودارها و آزمون دیکی-فولر (آزمون ریشه واحد) برای داده‌های E و U به ترتیب با مقادیر $-2/148$ و $-2/598$ که منجر به رد فرضیه صفر نمی‌شوند، نشان می‌دهند که داده‌ها مانا^۲ نیستند و باید مانا شوند. شکل ۲ نمودار داده‌های E و U را بعد از تفاضل‌گیری در مقابل زمان نشان می‌دهد و به نظر می‌رسد که داده‌ها با اولین تفاضل‌گیری مانا شده‌اند. از طرف دیگر، براساس آزمون دیکی-فولر برای متغیرهای تفاضل‌گیری شده E و U به ترتیب با مقادیر $-3/2669$ و $-3/7325$ فرضیه صفر رد می‌شود و مانایی داده‌های تفاضل‌گیری شده تأیید می‌شوند.

1. Balakrishnan, N. & Scarpa, B.

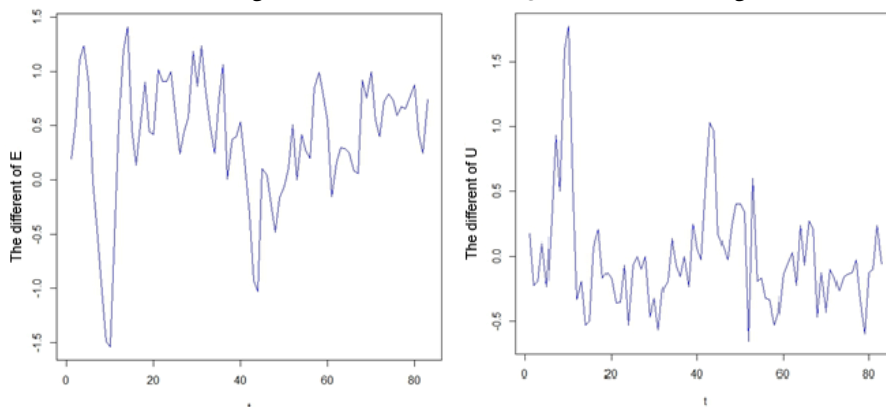
2. Stationary

شکل ۱. نمودار سری زمانی داده‌های کانادا برای متغیرهای E و U



مأخذ: یافته‌های پژوهش

شکل ۲. نمودار سری زمانی داده‌های کانادا بعد از تفاضل‌گیری E و U



مأخذ: یافته‌های پژوهش

براساس ملاک‌های آکائیکه، خان-کوئین و شوارتز-بیزی در جدول ۱ مرتبه مدل اتورگرسیو برداری، یک تعیین می‌شود. از طرف دیگر، با برازش مدل اتورگرسیو برداری مرتبه یک و محاسبه ریشه‌های معادله مشخصه که به ترتیب $1/215$ و $2/3917$ و بزرگتر از یک می‌باشند، شرط مانایی مدل تأیید می‌شود.

جدول ۱. مقادیر آکائیکه، حنان- کوئین و شوارتز- بیزی برای تعیین مرتبه مدل

مرتبه	۱	۲	۳	۴	۵
AIC	-۵/۱۷۱۹۴۵*	-۵/۱۶۷۱۵۷	-۵/۱۱۵۷۲۵	-۵/۱۵۱۵۴۵	-۵/۰۸۲۲۶۳
HQ	-۵/۰۹۶۶۹۲*	-۵/۰۴۲۱۱۸۴۴	-۴/۹۴۰۶۷۰	-۴/۹۲۶۴۷۴	-۴/۸۰۷۱۷۶
SC	-۴/۹۸۳۶۸۸*	-۴/۸۵۳۳۹۶	-۴/۶۷۶۴۵۹	-۴/۵۸۶۷۷۵	-۴/۳۹۱۹۸۸

علامت * مرتبه انتخاب شده توسط معیار را نشان می‌دهد.

مأخذ: یافته‌های پژوهش

براساس آزمون ماردیا، آزمون این که چولگی و کشیدگی با توزیع نرمال چندمتغیره مطابقت دارد یا خیر، مقدار چولگی ۰/۸۳۸۰۸۹ با مقدار p برابر ۰/۰۱۶۱۵۷ و مقدار کشیدگی برابر ۹/۳۳۹۹ با مقدار p برابر ۰/۹۹۷ در سطح ۵ درصد چولگی معنی‌دار است و کشیدگی معنی‌دار نیست بنابراین توزیع نرمال چوله چندمتغیره برای شوک‌ها در نظر گرفته می‌شود. بنابراین مدل اتورگرسو برداری مرتبه ۱ با فرض توزیع نرمال چندمتغیره و توزیع نرمال چوله چندمتغیره برای شوک‌ها در جدول ۲ برازش داده شده است.

جدول ۲. پارامترهای برآورد شده در مدل VAR(1) برای داده‌های تفاضل‌گیری شده کادانا

برآوردها	$\varepsilon_t \sim MN_2$	$\varepsilon_t \sim MSN_2$
$\hat{a}_{11}$	۰/۹۱۰۳	۰/۷۰۹۴
$\hat{a}_{12}$	۰/۲۱۳۹	-۰/۰۱۴۹
$\hat{a}_{21}$	-۰/۲۰۱۸	-۰/۴۶۶۳
$\hat{a}_{22}$	۰/۳۳۰۳	۰/۰۲۵۵
$\hat{s}_1$	-	۰/۰۷۰۳
$\hat{s}_2$	-	۰/۱۰۶۶
$\hat{\sigma}_{11}$	۰/۱۶۴۶	۰/۱۵۹۵
$\hat{\sigma}_{12}$	-۰/۰۹۱۰۶	-۰/۱۰۰۳
$\hat{\sigma}_{21}$	-۰/۰۹۱۰۶	-۰/۱۰۰۳
$\hat{\sigma}_{22}$	۰/۱۱۴۰۰	۰/۰۹۷۷

مأخذ: یافته‌های پژوهش

معیارهای AIC و BIC که در جدول ۳ آمده‌اند، نشان می‌دهند که برای این داده‌ها مدل اتورگرسو مرتبه ۱ با فرض توزیع نرمال چوله چندمتغیره برای شوک‌ها مناسب‌تر و کاراتر

از مدل اتورگرسیو مرتبه ۱ با فرض توزیع نرمال چندمتغیره برای شوک‌ها می‌باشد زیرا کمترین مقادیر آکائیکه، بیزی و منفی لگاریتم درست‌نمایی را نسبت به مدل دیگر دارد.

جدول ۳. معیارهای AIC و BIC برای مدل‌ها

توزیع شوک‌ها	AIC	BIC	-Log-Like
$\varepsilon_t \sim MN_2$	۲۸۲/۰۲۲۶	۲۸۶/۸۳۶	۱۳۹/۰۱۱
$\varepsilon_t \sim MSN_2$	۲۶۰/۶۷۸۱	۲۶۵/۴۹۱۵	۱۲۸/۳۳۹

مأخذ: یافته‌های پژوهش

بنابراین مدل برازش داده شده به صورت زیر نوشته می‌شود:

$$y_{1t} = 0.7094 y_{1t-1} - 0.0149 y_{2t-1} + \varepsilon_{1t}$$

$$y_{2t} = -0.4663 y_{1t-1} + 0.0255 y_{2t-1} + \varepsilon_{2t}$$

که در آن y_1 و y_2 به ترتیب متغیرهای تفاضل‌گیری شده E و U هستند.

۴-۲. داده‌های کاربردی ایران

در این بخش، متغیر AFE (شاخص کشاورزی، جنگل‌داری و ماهیگیری)^۱ را که درصدی از تولید ناخالص داخلی^۲ (GDP) است را در نظر می‌گیریم. صنعت کشاورزی مطابق با طبقه‌بندی استاندارد بین‌المللی صنعتی^۳ (ISIC) شامل چندین بخش کشاورزی، جنگل‌داری و ماهیگیری می‌شود که در تولید ناخالص داخلی اثرگذار است. به نظر می‌رسد این متغیر بر اشتغال زنان^۴ (EW) در صنعت تأثیرگذار است زیرا زنان نقش مهمی را در بخش‌های کشاورزی، شیلات و صنعت ایفا می‌کنند و پژوهش‌هایی در این زمینه‌ها توسط محققان صورت گرفته است که می‌توان به فریار و همکاران^۵ (۲۰۲۲) و اسفنجیر و رضایی^۶ (۱۳۹۹) اشاره نمود. داده‌های سالانه AFF و EW از سایت بانک جهانی^۷ برای سال‌های ۱۹۹۱ تا ۲۰۲۱ گرفته شده است. آزمون دیکی-فولر برای داده‌های AFE و EW به ترتیب با مقادیر

-
1. Agriculture, Forestry and Fishing
 2. Gross Domestic Product
 3. International Standard Industrial Classification
 4. Employment of Women
 5. Faryaar, H., et al.
 6. Esfanzir, A.A. & Rezaei R.H.
 7. <https://data.worldbank.org/country/iran-islamic-rep>

۰/۳۳۵۶ و ۰/۱۴۳۸ که منجر به رد فرضیه صفر نمی‌شوند، نشان می‌دهند که داده‌ها مانا نیستند و باید مانا شوند. شکل ۳ نمودار داده‌های AFF و EW را نسبت به زمان بعد از تفاضل‌گیری نشان می‌دهد و به نظر می‌رسد که داده‌ها با اولین تفاضل‌گیری مانا شده‌اند. از طرف دیگر، براساس آزمون دیکی-فولر برای متغیرهای تفاضل‌گیری شده AFE و EW به ترتیب با مقادیر ۳/۶۷۳۳- و ۳/۱۱۶۷- با مقادیر p -کمتر از ۵ درصد فرضیه صفر رد می‌شود و مانایی داده‌های تفاضل‌گیری شده تأیید می‌شوند. براساس ملاک‌های آکائیکه، حنان-کوئین و شوارتز-بیزی در جدول ۴ مرتبه مدل اتورگرسیو، ۱ در نظر گرفته می‌شود. همچنین با برازش مدل اتورگرسیو برداری مرتبه ۱ و محاسبه ریشه‌های معادله مشخصه که به ترتیب ۵/۸ و ۲۲/۱ و بزرگتر از ۱ می‌باشند، شرط مانایی مدل تأیید می‌شود.

جدول ۴. مقادیر آکائیکه، حنان-کوئین و شوارتز* بیزی برای تعیین مرتبه مدل

مدل	۱	۲	۳	۴	۵
AIC	-۰/۴۷۲۶۱ *	-۰/۴۱۰۹۳	-۰/۳۶۳۰۵۴	-۰/۴۲۶۱۷۵	-۰/۱۶۰۰۸
HQ	-۰/۳۹۴۴۸*	-۰/۲۸۰۷۰	-۰/۱۸۰۷۴۰	-۰/۱۹۱۷۷۱	-۰/۱۲۶۴۱۱
SC	-۰/۱۷۸۱۰*	-۰/۰۷۹۹۲۳	۰/۳۲۴۱۴	۰/۴۵۷۳۶۵	۰/۹۸۷۴۶۱

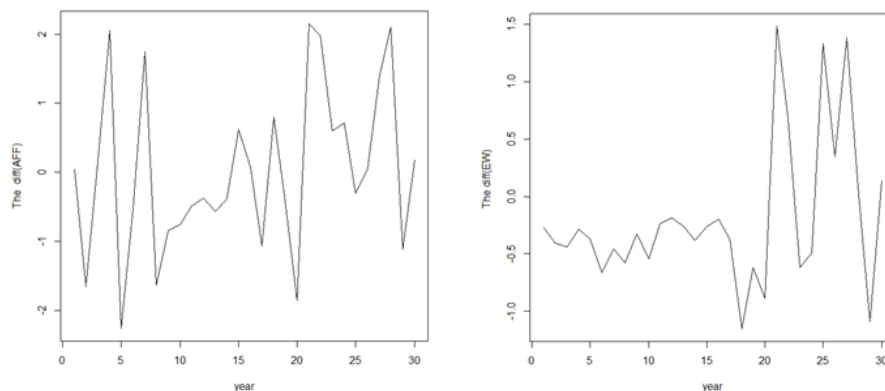
علامت * مرتبه انتخاب شده توسط معیار را نشان می‌دهد.

مأخذ: یافته‌های پژوهش

براساس آزمون Mardia، مقدار چولگی شوک‌ها برابر ۲/۲۵۰۰۴ با مقدار p -برابر ۰/۰۱۱۹ و مقدار کشیدگی برابر ۸/۸۹۴۱ با مقدار p -برابر ۰/۹۹۴۸ در سطح ۵ درصد نشان می‌دهد چولگی شوک‌ها نسبت به توزیع نرمال چندمتغیره معنی‌دار است و چولگی شوک‌ها تأیید می‌شود ولی کشیدگی معنی‌دار نیست بنابراین توزیع نرمال چوله چندمتغیره را برای شوک‌ها در نظر گرفته و مدل را برازش می‌دهیم.

برآورد پارامترهای مدل اتورگرسیو مرتبه اول با فرض توزیع نرمال چندمتغیره و توزیع نرمال چوله برای شوک‌ها و معیارهای AIC و BIC برای ارزیابی مدل‌ها در جدول ۵ آمده است. براساس آن‌ها می‌توان نتیجه گرفت که برای این داده‌ها مدل اتورگرسیو مرتبه ۱ با فرض توزیع نرمال چوله چندمتغیره برای شوک‌ها مناسب‌تر و کاراتر از مدل اتورگرسیو مرتبه ۱ با فرض توزیع نرمال چندمتغیره برای شوک‌ها است زیرا کمترین مقادیر آکائیکه، بیزی و منفی لگاریتم درست‌نمایی را نسبت به مدل دیگر دارد.

شکل ۳. نمودار سری زمانی داده‌های ایران بعد از تفاضل‌گیری AFF و EW



مأخذ: یافته‌های پژوهش

جدول ۵. پارامترهای برآورد شده در مدل VAR(1) برای داده‌های تفاضل‌گیری شده ایران

برآوردها	$\varepsilon_t \sim MN_2$	$\varepsilon_t \sim MSN_2$
$\hat{s}_1$	-	۰/۱۱۰۴
$\hat{s}_2$	-	-۰/۰۷۹۸
$\hat{\sigma}_{11}$	۱/۲۳۹۷	۱/۲۲۶۷
$\hat{\sigma}_{12}$	۰/۲۷۴۸	۰/۲۸۳۶
$\hat{\sigma}_{21}$	۰/۲۷۴۸	۰/۲۸۳۶
$\hat{\sigma}_{22}$	۰/۳۸۵۵	۰/۳۷۹۶
AIC	۱۴۶/۱۱۴۵	۱۴۵/۹۱۸۵
BIC	۱۴۸/۸۴۹۱	۱۴۸/۶۵۳۱

مأخذ: یافته‌های پژوهش

بنابراین مدل برازش داده شده می‌تواند به صورت زیر نوشته شود:

$$y_{1t} = -0.25626 y_{1t-1} - 0.78496 y_{2t-1} + \varepsilon_{1t}$$

$$y_{2t} = -0.09239 y_{1t-1} + 0.29052 y_{2t-1} + \varepsilon_{2t}$$

که در آن y_1 و y_2 به ترتیب متغیرهای تفاضل‌گیری شده AFF و EW هستند.

۵. نتیجه‌گیری

تحلیل سیاست‌گذاری‌های کلان اقتصادی مستلزم توجه به بررسی و تحلیل صحیح داده‌ها و مدل‌سازی مناسب برای آن‌ها است. هدف این مقاله گسترش مدل سری زمانی چندمتغیره بوده است زیرا این مدل‌ها در پژوهش‌های داخلی و خارجی در زمینه‌های مالی و اقتصادی نقش مهمی را ایفا می‌کنند. نوآوری این مقاله، در نظر گرفتن مدل اتورگرسو چندمتغیره مرتبه p ($VAR(p)$)، با توزیع نرمال چوله چندمتغیره برای شوک‌ها است زیرا در برخی موارد در مسائل واقعی مالی، اقتصادی و پزشکی شوک‌های مدل از توزیع نرمال چندمتغیره پیروی نمی‌کنند. برای برآورد پارامترهای این مدل از روش بیشینه درست‌نمایی استفاده شد. از آن‌جاکه فرم‌های بسته برای برآورد پارامترها از حل مشتقات حاصل نمی‌شود، الگوریتم ECM برای برآورد پارامترهای مدل به کار گرفته شد. سپس با استفاده از دو مجموعه داده‌های واقعی کشور کانادا (برای مدل‌بندی هم‌زمان متغیرهای E و U) و کشور ایران (برای مدل‌بندی هم‌زمان متغیرهای AFF و EW) که شوک‌ها دارای چولگی بوده‌اند، مدل VAR با مرتبه تعیین شده برای داده‌ها با در نظر گرفتن توزیع نرمال چوله چندمتغیره و توزیع نرمال چندمتغیره برای شوک‌ها برازش داده و مشاهده شد که در این داده‌ها برای شوک‌ها، مدل اتورگرسو برداری با توزیع نرمال چوله چندمتغیره طبق معیارهای اطلاعاتی AIC و BIC مناسب‌تر و کاراتر از مدل با توزیع نرمال چندمتغیره بوده است.

تعارض منافع

تعارض منافع وجود ندارد.

ORCID

Manijeh Mahmoodi
Mohammad Reza Salehi Rad



<https://orcid.org/0009-0003-7143-2396>
<http://orcid.org/0000-0003-1519-7590>

منابع

اسفنجیر عباسی، علی‌اصغر و رضایی روشن، هدی. (۱۳۹۹). بررسی تأثیر اشتغال زنان بر رشد اقتصادی در کشورهای منتخب خاورمیانه. فصلنامه علمی پژوهشی زن و جامعه، ۲(۴۲)، ۲۰۷-۲۲۶.

- حیدری، حسن. (۱۳۹۰). مدل VAR جایگزین برای پیش‌بینی تورم ایران: کاربردی از تبدیل
Bewley. *پژوهش‌های اقتصادی ایران*، ۱۶ (۴۶)، ۷۷-۹۶.
- خوهرسندی، مرتضی، محمدی، تیمور، ارباب، حمیدرضا و سخایی، عمادالدین. (۱۴۰۱). آثار
شوگ‌های اقتصادی خارجی بر متغیرهای کلان اقتصادی ایران: رویکرد خودرگرسیون
برداري جهانی (GVAR). *پژوهش‌های اقتصادی ایران*، ۲۷ (۹۱)، ۹-۵۰.
<http://dx.doi.org/10.22054/ijer.2020.52537.868>
- محمدی، تیمور، عزیزخانی، فاطمه، طایی، حسن و بهرامی، جاوید. (۱۳۹۸). پویایی‌های کلان
اقتصادی مقررات‌زدایی در بازارهای محصول و کار در کشورهای منا: رهیافت PANEL
VAR. *پژوهش‌های اقتصادی ایران*، ۲۴ (۸۰)، ۳۷-۶۷.
<https://doi.org/10.22054/ijer.2019.11112>

References

- Akaik, H. (1974). A new look at the statistical model identification. *IEEE Trans Automat Contr.* 19, 716-723.
doi.org/10.1109/TAC.1974.1100705
- Arellano-Valle, R. B., Ozan, S., Bolfarine, H., Lachos, V.H. (2005). Skew-normal measurement error models. *Journal of Multivariate Analysis*, 96, 265-281. doi.org/10.1016/j.jmva.2004.11.002
- Arellano-Valle, R.B. & Azzalini, A. (2006). On the unification of families of skew-normal distributions. *Scandinavian Journal of Statistics*, 33, 561-574. doi.org/10.1111/j.1467-9469.2006.00503.x
- Arellano-Valle, R.B. & Genton, M.G. (2005). Fundamental skew distributions. *Journal of Multivariate Analysis*, 96, 93-116.
doi.org/10.1016/j.jmva.2004.10.002
- Arellano-Valle, R.B., Gomez, H.W. & Quintana, F.A. (2004). A new class of skew-normal distributions. *Communication in Statistics- Theory and Methods*, 33, 1465-1480. doi.org/10.1081/STA-120037254
- Azzalini, A. & Capitanio, A. (1999). Statistical application of the multivariate skew-normal distribution. *Journal of Royal Statistical Society, series B*, 61, 579-602. doi.org/10.1111/1467-9868.00194
- Azzalini, A. & Capitanio, A. (2014). *The Skew-Normal and Related Families*. Cambridge CB2 8BS, United Kingdom.
- Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 171-178.
doi.org/10.6092/ASSN.1973-2201/711
- Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *Scandinavi Journal of Statistics*, 32, 159-188.
doi.org/10.1111/1467-9469.2005.00426.x
- Balakrishnan, N. & Scarpa, B. (2012). Multivariate measures of skewness for the skew-normal distribution. *Journal of Multivariate Analysis*, 104, 73-87. [doi.org 10.1016/j.jmva.2011.06.017](https://doi.org/10.1016/j.jmva.2011.06.017)

- Bondon, P. (2009). Estimation of autoregressive models with epsilon-skew-normal innovations. *Journal of Multivariate Analysis*, 100(8), 1761-1776. doi.org/10.1016/j.jmva.variate2009.02.006
- Box, G. & Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francis.
- Esfanjir, A.A. & Rezaei R.H. (1399). Investigating the effect of women's employment on economic growth in selected countries of Middle East. *Quarterly Journal of Woman and Society*. 11,207-226. [In Persian]
- Faryaar, H., Macdonald, R. & Watt, J. (2022). Improving the measurement of the contribution of women to the economy: estimates of GDP. *Economic analysis division at statistics Canada*.36, 28-001. doi.org/10.25318/36280001202201000003-eng
- Ghasami, S., Khodadadi, Z. & Maleki, M. (2019). Autoregressive processes with generalized hyperbolic innovations. *Communication in Statistics-Simulation and Computation*.49(12), 3080-3092. doi.org /10.1080/03610918.2018. 1535066
- Gupta, A. & Chang, C. (2002). Multivariate skew-symmetric distributions. *Applied Mathematics Letters*, 16(5), 634-646. doi.org/10.1016/S0893-9659(03)00060-0
- Heidari, H. (2011). An alternative VAR model for forecasting Iranian inflation: An application of bewley transformation. *Iranian Journal of Economic Research*, 46, 77-96. [In Persian]
- Khorsandi, M., Mohammadi, T., Arab, H. & Sakhaei, E. (2022). The effect of external economic shocks on Iran's macroeconomic variable: Global VAR approach. *Iranian Journal of Economic Research*, 27, 9-50. doi.org/10.22054/ijer.2020.52537.868 [In Persian]
- Kilian, L. & Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*. (Themes in Modern Econometrics), Cambridge, United Kingdom.
- Koop, G. & Korobilis, D. (2009). Bayesian multivariate time series methods for empirical macroeconomics. *Foundation and Trends in Econometrics*, 3, 267-351. doi.org/10.1561/08000000013
- Liu, J., Kumar, S. & Palomar, D. (2019). Parameter estimation of heavy-tailed AR model with missing data via stochastic EM. *IEEE Transaction on Signal Processing*, 67(8), 2159- 2172. doi.org/10.1109/ TSP.2019. 2899816
- Liu, Y., Sang, R. & Liu, Sh. (2016). Diagnostic analysis for a vector autoregressive model under Student's t- distributions statistic. *Neerlandica*, 71(2), 86- 114. doi.org/ 10.1111/stan.12102
- Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Springer-Verlag. Berlin and New York.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer Science & Business Media.
- Lütkepohl, H. (2020). Structural vector autoregressive models with more shocks than Variables identified via heteroscedasticity. *Economics Letters*, 195, 19510458. doi.org /10.1016/j.econlet.2020.109458

- Maleki, M., Wraith, D., Mahmoudi, R. & Javier, E. (2019). Asymmetric heavy-tailed vector auto-regressive processes with application to financial data. *Journal of Statistical Computation and Simulation*, 90(2), 324-340. doi.org/10.1080/00949655.2019.168067
- Mardia, K. (1970). Measure of multivariate skewness and kurtosis with application. *Biometrika*, 36, 519-530. doi.org/10.1093/biomet/57.3.519
- Mardia, K. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhyā*, ser. B. 36, 115-128. doi.org/10.1080/00369791.1974.11659939
- Mirzaei, H., Razban, N., Mohammadi, T. & Morovat, H. (2023). Analyzing the housing market network among Iran's Provinces: New evidence through variance decomposition of forecast errors, 88, 120-157. doi.org/10.22054/joer.2024.75890.1159
- Mohammadi, T., Azizkhani, F., Taei, H. & Javid, B. (1398). Macroeconomic dynamics of deregulation in product markets and work in Mena countries: Panel VAR. *Iranian Journal of Economic Research*, 80, 37-67. doi.org/10.22054/ijer.2019.11112
- Neethling, A., Ferreira, J. & Naderi, M. (2020). Skew generalized normal innovations for AR(p) process endorsing asymmetry. *Symmetry*, 12(8), 1253. doi.org/10.3390/sym12081253
- Ni, Sh. & Sun, D. (2005). Bayesian estimates for Vector autoregressive. *Journal of Business and Economic Statistics*, 23, 105-117. doi.org/10.1198/073500104000000622.
- Pourahmadi, M. (2001). *Foundation of Time Series Analysis and Prediction Theory*; John Wiley & Sons, Inc: Hoboken, NJ, USA.
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann Stat*, 6(2), 461-464. doi.org/10.1214/aos/1176344136
- Sharafi, M. & Nematollahi, A. (2016). AR(1) model with skew-normal innovations. *Metrika*, 79, 1011-1029. doi.org/10.1007/s00184-016-0587-7.
- Sims, C. (1980). Macroeconomic and reality. *Econometric*, 48, 1-48. doi.org/10.2307/1912017
- Stock, J.H. & Watson, M.W. (2001). Vector autoregressive. *Journal of Economic Perspectives*, 15(4), 101-115. doi.org/10.1257/jep.15.4.10
- Sujit, S., Dipak, D. & Marco, B. (2003). A new class of multivariate distributions with applications to Bayesian regression models. *Canadian Journal of Statistics*, 129(31), 129-150. doi.org/10.2307/3316064
- Tarami, B. & Pourahmadi, M. (2003). Multi-variate t autoregressive: innovations, prediction variances and exact Likelihood equations. *Journal of Time Series Analysis*, 24, 739-754. doi.org/10.1111/j.1467-9892.2003.00332.x
- Tsay, R.S. (2002). *Analysis of Financial Time Series*. John Wiley & Sons, Inc. ISBN: 0-471-41544-8

- Tsung, L. (2009). Maximum likelihood estimation for multivariate skew normal mixture models. *Journal of Multivariate Analysis*, 100(2), 257-265. doi.org/10.1016/j.jmva.2008.04.010
- Tsung, L., Hsiu, H. & Chiang, Chen. (2009). Analysis of multivariate skew normal models with incomplete data. *Journal of Multivariate Analysis*, 100, 2337-2351. doi.org/10.1016/j.jmva.2009.07.005
- Wai, C., Mumtaz, H. & Pinter, G. (2017). Forecasting with VAR models: fat tails and stochastic volatility. *International Journal of Forecasting*, 33(4), 1124-1143. doi.org/ 10.1016/j.ijforecast.2017.03.001
- Wei, W. (2006). *Time Series Analysis Univariate and Multivariate Methods*. Boston, Pearson Addison Wesley.
- Zamani Mehryan, S. & Sayyareh, A. (2015). Statistical inference in autoregressive models with non-negative residuals. *Statistical Research and Training Center*, Iran JSRI 2015. 12(1), 83-104.

استناد به این مقاله: محمودی، منیژه و صالحی‌راد و محمدرضا. (۱۴۰۴). برآورد مدل اتورگرسیون برداری با توزیع نرمال چوله چندمتغیره برای شوک‌ها و تحلیل عملکرد آن برای دو مجموعه داده واقعی. *پژوهش‌های اقتصادی ایران*، ۳۰(۱۰۲)، ۳۸-۶۳.



Iranian Journal of Economic Research is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.