International Journal for Innovative Engineering and Management Research

PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL

www.ijiemr.org

COPY RIGHT

ELSEVIER SSRN

IJIEMR Transactions, online available on 30th Nov 2020. Link :http://www.ijiemr.org/downloads.php?vol=Volume-09&issue=Issue 11

10.48047/IJIEMR/V09/ISSUE 11/55

Title **Parameter Estimation for a Multivariate Time Series VAR Model**

Volume 09, ISSUE 11, Pages: 297-305

Paper Authors **Dr.M. Chinna Giddaiah**

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Parameter Estimation for a Multivariate Time Series VAR Model

Dr.M. Chinna Giddaiah
Lecturer in Statistics,
VRS Degree College, Veerapunayunipalli, YSR Kadapa. India
Mail ID: mcgvrsdc@gmail.com

Abstract

Building a multivariate time series model involves five pivotal steps: Identification Specification, Estimation and Hypothesis Testing, Diagnostic Assessment, and Forecasting. Estimating parameters in a multivariate Vector Autoregressive (VAR) model presents a greater challenge compared to univariate autoregressive models. Under the assumption of normality in error distributions, Maximum Likelihood Estimation (MLE) and the Likelihood Ratio test are applicable in the context of multivariate VAR models. In this research article, we embark on a journey to estimate the parameters of a multivariate VAR model. We employ the method of Maximum Likelihood Estimation based on ordinary least squares regression. To enhance the accuracy of our model, we estimate the dispersion matrix of errors using Internally Studentized residuals. Furthermore, we introduce a test procedure for determining the optimal number of lags for variables within the multivariate VAR model, leveraging the power of the Likelihood Ratio test.

INTRODUCTION: -

Univariate time series models offer the advantage of predicting a variable solely based on its past, present, and future values. However, these univariate models can significantly enhance their explanatory power by incorporating political-economic information contained in interacting variables.

In contrast, a Multivariate Time Series Model serves as a versatile, unrestricted approximation to the reduced form of an unknown structural specification within a simultaneous equations model. Pioneering work by Zellner and Palm in 1974 and subsequent research by Zellner in 1979 have demonstrated that any structural model can be reformulated in the shape of a multivariate time series model.

The field of time series analysis encompasses both linear and nonlinear methods and spans the realms of univariate and multivariate approaches. In disciplines such as statistics, econometrics, quantitative finance, seismology, meteorology, and geophysics, the primary objective of time series analysis revolves around "forecasting."

Conversely, in contexts such as data mining, pattern recognition, and machine learning, time series analysis finds applications beyond forecasting. It extends to tasks like clustering, classification, content-based querying, anomaly detection, as well as predictive modeling. Time series analysis thus emerges as a pivotal tool for exploring data dynamics and extracting valuable insights across diverse domains and applications.

Vector Autoregressive (VAR) Models: Unraveling Multivariate Time Series Analysis

Vector Autoregressive (VAR) models are a class of multivariate time series models extensively used in various fields, such as economics, finance, econometrics, and data science. These models play a crucial role in capturing and understanding the dynamic relationships between multiple time series variables.

Fundamental Concepts:

1. **Multivariate Time Series Data:** VAR models are designed for multivariate time series data, which involves a collection of related variables observed at multiple time points. These variables can be economic indicators, financial variables, or any other dataset where the interactions between variables are of interest.
2. **Lagged Relationships:** VAR models consider the past values of all variables in the system to predict their future values. This feature makes VAR models different from univariate models, where each variable is predicted based on its own historical values.
3. **Simultaneity:** VAR models acknowledge that variables within the system can influence each other simultaneously. This captures the complex interdependencies often observed in real-world data.

Model Structure:

A VAR model can be represented mathematically as follows:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2+\dots} \dots \dots \dots + A_P Y_{t-P} + \varepsilon_t$$

Here:

- Y_tis a vector of multivariate time series at time t
- A_1, A_2, ..., A_p are coefficient matrices that relate past values of Y to the current value.
- prepresents the order of the VAR model, indicating how many lags are considered.
- ε_t is the white noise error term.

Applications:

VAR models are applied in various areas:

1. **Macroeconomics:** They are used to study the dynamic relationships among economic variables like GDP, inflation, and unemployment.
2. **Finance:** VAR models help in analyzing the joint behavior of financial market variables, such as stock prices, interest rates, and exchange rates.
3. **Forecasting:** VAR models are valuable for short- and medium-term forecasting of multivariate time series data.
4. **Policy Analysis:** They are employed to assess the impact of economic policies on different variables within an economy.

5. **Risk Management:** In financial risk management, VAR models are useful for quantifying and understanding the risk associated with portfolios.

Model Estimation:

Estimating a VAR model involves determining the coefficients ($A_1, A_2, ..., A_p$) and the order of the model ($p). The most common estimation techniques include ordinary least squares (OLS) and maximum likelihood estimation (MLE).

Impulse Response Analysis:

One key advantage of VAR models is the ability to perform impulse response analysis. This technique helps assess how a shock to one variable affects all the variables in the system over time, shedding light on dynamic relationships.

In summary, Vector Autoregressive (VAR) models are a powerful tool for understanding and forecasting multivariate time series data. They allow researchers and analysts to capture the intricacies of interactions between multiple variables, making them invaluable in various domains, from economics and finance to data science and policy analysis.

Maximum Likelihood Estimates (MLE) of VAR Parameters using OLS Regression

Vector Autoregressive (VAR) models are typically estimated using the MLE method, which maximizes the likelihood function of the observed data given the model. However, when the assumptions of the MLE method cannot be met, or in practical situations where MLE is challenging, an alternative approach is to estimate VAR parameters using OLS regression. Here's how this can be done:

Step 1: Model Specification

- Start by specifying the VAR model order (p), which represents the number of lagged observations you want to include. For simplicity, let's consider a VAR(1) model as an example.
- Define your multivariate time series data, represented as Y_t, where t denotes time.
- The VAR(1) model is given by: $Y_t = A_1 Y_{(t-1)} + \varepsilon_t$, where A_1 is the coefficient matrix for lag 1 and ε_t is the error term at time t.

Step 2: Data Preparation

- Prepare your data. Ensure it is stationary, or make it stationary through differencing, if necessary, to meet the OLS assumptions.

Step 3: OLS Estimation

- For a VAR(1) model, you can estimate the coefficients A_1 using OLS regression. Each equation in the VAR system will have a set of coefficients to estimate.
- Set up the OLS regression for each equation. The equation for the i-th variable in the VAR(1) model is as follows: $Y_{i,t} = a_i + B_i Y_{i,t-1} + \varepsilon_{i,t}$, where a_i is the intercept, B_i is the coefficient for lag 1, and $\varepsilon_{i,t}$ is the error term.
- Estimate the coefficients for each equation using standard OLS regression. This involves minimizing the sum of squared residuals by varying the

coefficients. In matrix form, this can be represented as follows: $Y = X\beta + \varepsilon$, where Y is the vector of observed data, X is the matrix of lagged values, β is the coefficient matrix to be estimated, and ε is the error term.

- Use standard OLS software or libraries to perform the estimation. In practice, statistical software like Python, R, or dedicated econometrics software can be used for this purpose.

Step 4: MLE for Residuals

- After estimating the coefficients using OLS, you can use the residuals from the OLS regression as the input to calculate the MLE of the error covariance matrix (Σ).
- The MLE for Σ can be obtained by maximizing the likelihood function based on the residual vector ε_t. This step ensures that the error structure of the model follows a multivariate normal distribution.

Step 5: Hypothesis Testing

- Once the MLE is obtained, you can perform hypothesis tests on the VAR parameters or conduct further analysis as needed, such as impulse response analysis, Granger causality, or forecasting.

It's important to note that while OLS is a simpler and more accessible method for VAR model estimation, MLE is generally considered a more appropriate estimation technique when the assumptions of the model are satisfied. However, in situations where MLE may not be feasible or when a quick approximation is needed, OLS estimation can be a viable alternative.

ESTIMATING Φ BY USING STUDENTIZED RESIDUALS:

Estimating the parameter φ (phi) using studentized residuals typically involves a statistical procedure to assess the significance of the parameter in a model, such as a time series model or regression model. Studentized residuals are adjusted residuals that take into account the uncertainty in the estimated model parameters. The parameter φ is often associated with autoregressive (AR) or moving average (MA) terms in time series models.

Here's a general outline of how to estimate the parameter φ using studentized residuals:

1. Model Specification:

- Begin by specifying the time series model or regression model that contains the parameter φ. For example, in an autoregressive model (AR), φ represents the autoregressive coefficient.

2. Model Estimation:

- Estimate the model parameters, including φ, using the chosen estimation method. Common methods include maximum likelihood estimation (MLE) for time series models or ordinary least squares (OLS) for regression models.

3. Calculate Residuals:

- Compute the model residuals by subtracting the predicted values from the observed data.

4. Calculate Studentized Residuals:

- Studentized residuals are calculated by dividing the model residuals by their estimated standard errors. This standardization allows for the comparison of residuals across different models and datasets, as it accounts for variations in the residual variance.

5. Perform a Hypothesis Test:

- To estimate the parameter φ, you can perform a hypothesis test. This test typically involves a null hypothesis that φ equals a specific value (e.g., φ = 0), indicating no effect, versus an alternative hypothesis that φ is not equal to that value.

6. Test Statistic Calculation:

- Calculate the test statistic using the studentized residuals and the null hypothesis. The specific test statistic depends on the chosen hypothesis test. For example, a t-test statistic is commonly used for testing the significance of φ in regression models.

7. Significance Testing:

- Determine the statistical significance of the parameter φ by comparing the calculated test statistic to a critical value from a probability distribution (e.g., t-distribution). The critical value depends on the chosen significance level (e.g., 0.05 for a 5% significance level).

8. Estimation of φ:

- If the test statistic is statistically significant (i.e., it exceeds the critical value), you can conclude that the parameter φ is not equal to the null hypothesis value. In this case, you can estimate the value of φ based on the test statistic and standard errors.

It's important to choose an appropriate hypothesis test and significance level based on the specific research question and context. The estimation of φ using studentized residuals allows you to assess the statistical significance of the parameter within the chosen model. The procedure may vary depending on the type of model and hypothesis test being conducted.

TESTING NUMBER OF LAGS OF VARIABLE FOR VAR MODEL BY USING THE LIKELIHOOD RATIO TEST

Testing the number of lags in a Vector Autoregressive (VAR) model using the Likelihood Ratio Test is a common approach to determine the appropriate lag order for the model. This test helps find the optimal number of lags that best fits the data and minimizes unnecessary complexity. Here's how you can perform the Likelihood Ratio Test for lag order selection in a VAR model:

1. Model Specification:

- Begin by specifying the VAR model, including the maximum lag order you want to consider. A VAR(p) model includes p lagged values of the variables.

2. Model Estimation:

- Estimate multiple VAR models with different lag orders, ranging from a minimum lag order (usually 1) to the maximum lag order you have specified.

3. Likelihood Function:

- Calculate the likelihood function for each estimated VAR model. The likelihood function measures how well the model fits the data.

4. Restricted and Unrestricted Models:

- Create two versions of each VAR model: the "unrestricted" model with the maximum lag order and a "restricted" model with a reduced lag order (e.g., one lag less).

5. Likelihood Ratio Statistic:

- Calculate the likelihood ratio statistic, denoted as LR, by comparing the likelihood of the restricted model (with fewer lags) to the likelihood of the unrestricted model (with more lags). The likelihood ratio is calculated as follows:

LR = -2 * [log likelihood of the restricted model - log likelihood of the unrestricted model]

6. Degrees of Freedom:

- Determine the degrees of freedom for the likelihood ratio test. It is equal to the difference in the number of parameters between the restricted and unrestricted models. The degrees of freedom are usually (p - q), where p is the number of lags in the unrestricted model, and q is the number of lags in the restricted model.

7. Null Hypothesis:

- Formulate the null hypothesis (H0) that the additional lags in the unrestricted model do not provide a statistically significant improvement in model fit compared to the restricted model.

8. Critical Value:

- Choose a significance level (e.g., 0.05) and find the critical value from a chi-squared distribution table or calculator based on the chosen significance level and degrees of freedom.

9. Likelihood Ratio Test Statistic Comparison:

- Compare the calculated LR statistic to the critical value. If the LR statistic exceeds the critical value, reject the null hypothesis, indicating that the additional lags in the unrestricted model significantly improve the model fit.

10. Optimal Lag Order:

- The optimal lag order is determined based on the largest lag order for which the null hypothesis is not rejected. This lag order is considered the most appropriate for the VAR model.

By following these steps and conducting the Likelihood Ratio Test for different lag orders, you can select the lag order that strikes a balance between model complexity and goodness of fit, ultimately providing you with an appropriate VAR model for your data.

Algorithm :-

To test for the optimal number of lags in a Vector Autoregression (VAR) model using the likelihood ratio test, you can follow these steps. The likelihood ratio test is used to compare two nested models: one with k lags and another with k-1 lags.

Step 1: Define the VAR Model

- Define your VAR(p) model, where 'p' is the number of lags.
- Specify the variables and their order.
- Decide the maximum number of lags to consider (p_max).

Step 2: Estimate the VAR(p_max) Model

- Estimate the VAR model with p_max lags.
- This will serve as the unrestricted or larger model.
- You can use statistical software like R, Python (with libraries like statsmodels or VARMAX), or specialized time series software for this estimation.

Step 3: Estimate the VAR(p_max-1) Model

- Estimate the VAR model with p_max-1 lags.
- This will be the restricted or smaller model.

Step 4: Calculate Likelihood Ratios

- Compute the log-likelihood for both the VAR(p_max) and VAR(p_max-1) models.
- Calculate the likelihood ratio statistic (LR) as the difference in log-likelihoods between the larger and smaller models:
    - LR = 2 * (log-likelihood of VAR(p_max) - log-likelihood of VAR(p_max-1))

Step 5: Perform Likelihood Ratio Test

- Under the null hypothesis (H0), the smaller model (VAR(p_max-1)) is true.
- Under the alternative hypothesis (H1), the larger model (VAR(p_max)) is true.
- The LR statistic follows a chi-squared distribution with degrees of freedom equal to the difference in the number of parameters between the two models.

Step 6: Set Significance Level

- Choose a significance level (e.g., 0.05) to determine whether to reject the null hypothesis.

Step 7: Compare LR Statistic to Critical Value

- Calculate the critical value from the chi-squared distribution table for the chosen significance level and degrees of freedom.
- Compare the LR statistic to the critical value.
- If the LR statistic is greater than the critical value, you reject the null hypothesis (H0) in favor of the alternative hypothesis (H1), indicating that the VAR(p_max) model is preferred.
- If the LR statistic is less than the critical value, you fail to reject the null hypothesis, indicating that the VAR(p_max-1) model is preferred.

Step 8: Interpretation

- If you reject the null hypothesis, it suggests that the VAR(p_max) model is preferred, and you should proceed with p_max lags.
- If you fail to reject the null hypothesis, it suggests that the VAR(p_max-1) model is preferred, and you should choose p_max-1 lags.

Repeat these steps with different lag values and compare the results to select the optimal number of lags for your VAR model. The likelihood ratio test helps you find the number of lags that best captures the temporal dependencies in your time series data.

Conclusions :-

Time series data is characterized by a sequence of values, all measured on the same scale and indexed by a time-related parameter. These datasets can exhibit an astonishing array of patterns and shapes, reflecting the diverse underlying functions they represent. In fact, the number of potential time series is equivalent to the number of real-number functions. Concepts closely linked to time series encompass longitudinal data, growth curves, repeated measures, economic models, multivariate analysis, signal processing, and system analysis.

The parameters of a Vector Autoregression (VAR) model are typically estimated using the maximum likelihood estimation method based on ordinary least squares regression. Additionally, the dispersion matrix of errors in a VAR model can be estimated using internally studentized residuals, which helps in understanding and quantifying the uncertainty in the model.

Furthermore, a test procedure has been developed for assessing the appropriate number of lags for the variables in a VAR model. This procedure relies on internally studentized residuals to make informed decisions about the lag structure of the model, ensuring that it captures the underlying dynamics of the data effectively.

References :-

1. Abraham, B., and Johness, L. (2005) - The book "Statistical Methods for Forecasting" authored by Abraham and Johness provides a comprehensive overview of statistical techniques for forecasting.
2. Box, George E. P., Jenkins, Gwilym M., and Reinsel, Gregory C. (2015). "Time Series Analysis: Forecasting and Control." This influential book introduces Box-Jenkins methodology for time series analysis.
3. Brock, William A., and Durlauf, Steven N. (2001). "Econometrics: Theory and Applications with EViews." This book covers econometric theory and its application using the EViews software.
4. Brockwell, Peter J., and Davis Richard A. (2002) - "Introduction to Time Series and Forecasting" is a textbook that covers the fundamentals of time series analysis and forecasting.
5. Enders, Walter (2014). "Applied Econometric Time Series." A widely used textbook that covers applied econometrics and time series analysis with practical examples.
6. Fuller W.A (1996) - Fuller's book, "Introduction to Statistical Time Series," is a resource for

understanding statistical aspects of time series data, and it offers insights into modeling and analysis.

7. Granger, Clive W. J., and Newbold, Paul (1977). "Forecasting Economic Time Series." This classic work focuses on forecasting techniques for economic time series data.
8. Hamilton, James D. (1994). "Time Series Analysis." This comprehensive book provides an in-depth exploration of time series models, including both theory and practical applications.
9. Lütkepohl, Helmut (2007). "New Introduction to Multiple Time Series Analysis." An introduction to multivariate time series analysis and its applications.
10. Pankratz, Alan (1983). "Forecasting with Univariate Box-Jenkins Models: Concepts and Cases." A practical guide to univariate time series forecasting using Box-Jenkins models.
11. Pesaran, M. Hashem, and Pesaran, Bahram (2010). "Time Series Econometrics Using Microfit 5.0: A User's Manual." A practical guide to time series econometrics using the Microfit software.
12. Shumway, Robert H., and Stoffer, David S. (2016). "Time Series Analysis and Its Applications: With R Examples." This book combines theory and practical applications using the R programming language.
13. Sims, Christopher (Jan 1980) - Sims' paper, published in Econometrica, delves into the relationship between macroeconomics and reality, offering insights into the challenges and complexities of modeling real-world economic systems.
14. Tsay, Ruey S. (2005). "Analysis of Financial Time Series." Focusing on financial time series data, this book explores statistical methods and models for analyzing financial markets.