

grangersearch: An R Package for Granger Causality Testing with Tidyverse Integration

Nikolaos Korfiatis
Department of Informatics
School of Information Science and Informatics
Ionian University
Plateia Iatrou Tsirigoti 7, GR-49100
Corfu, Greece
nkorf@ionio.gr

January 2, 2026

Abstract

This paper introduces **grangersearch**, an R package for performing exhaustive Granger causality searches on multiple time series. The package provides a user-friendly interface for discovering predictive causal relationships between variables using Vector Autoregressive (VAR) models. Key features include: (1) exhaustive pairwise search across multiple variables, (2) automatic lag order optimization with visualization, (3) tidyverse-compatible syntax with pipe operators and non-standard evaluation, (4) built-in visualization for lag selection analysis, and (5) integration with the **broom** ecosystem through **tidy()** and **glance()** methods. The package is designed for researchers and practitioners in economics, finance, and other fields who need to discover temporal causal relationships among multiple time series. This paper describes the statistical methodology, demonstrates the package functionality through examples, and discusses practical considerations for applying Granger causality tests.

Keywords: Granger causality, time series analysis, VAR models, R package, tidyverse

1 Introduction

Understanding causal relationships between time series variables is a fundamental problem in economics, finance, neuroscience, and many other fields. While true causality is philosophically complex and difficult to establish from observational data alone, Granger (1969) proposed a practical, testable notion of causality based on predictability: a variable X is said to “Granger-cause” another variable Y if past values of X contain information that helps predict Y beyond what is contained in past values of Y alone.

Granger causality has become one of the most widely used methods for analyzing lead-lag relationships in time series data. Applications span diverse domains including macroeconomics (Sims, 1972), financial markets (Hiemstra and Jones, 1994), energy economics (Kraft and Kraft, 1978), and neuroscience (Seth et al., 2015).

Despite its popularity, implementing Granger causality tests in R requires users to navigate multiple packages and manually construct the testing procedure. The **vars** package (Pfaff, 2008) provides the underlying VAR modeling infrastructure, but users must understand the model specification, extract appropriate test statistics, and interpret results. This complexity can be a barrier for applied researchers who need a straightforward tool for exploratory causal analysis.

This paper presents **grangersearch**, an R package that provides a simple, intuitive interface for Granger causality testing. The package wraps the **vars** infrastructure while providing:

- A single function interface that tests causality in both directions
- Exhaustive pairwise search across multiple time series variables
- Automatic lag order optimization with visualization
- Tidiverse-compatible syntax with pipe operators and non-standard evaluation
- Structured output objects with all relevant test statistics
- Integration with the **broom** package ecosystem
- Comprehensive input validation and informative error messages

The remainder of this paper is organized as follows. Section 2 reviews the statistical methodology underlying Granger causality tests. Section 3 describes the package design and functionality. Section 4 provides worked examples. Section 5 discusses practical considerations and limitations. Section 6 concludes.

2 Statistical Methodology

2.1 Granger Causality

The concept of Granger causality is based on two principles: (1) the cause occurs before the effect, and (2) the cause contains unique information about the effect that is not available elsewhere (Granger, 1969).

Formally, let $\{X_t\}$ and $\{Y_t\}$ be two stationary time series. X is said to Granger-cause Y if:

$$\sigma^2(Y_t|Y_{t-1}, Y_{t-2}, \dots) > \sigma^2(Y_t|Y_{t-1}, Y_{t-2}, \dots, X_{t-1}, X_{t-2}, \dots) \quad (1)$$

where $\sigma^2(Y_t|\cdot)$ denotes the variance of the optimal linear predictor of Y_t given the conditioning information. In other words, X Granger-causes Y if including past values of X improves the prediction of Y beyond what is achieved using past values of Y alone.

2.2 Vector Autoregressive Models

The standard approach to testing Granger causality uses Vector Autoregressive (VAR) models (Sims, 1980). A VAR model of order p for two variables can be written as:

$$Y_t = \alpha_1 + \sum_{i=1}^p \beta_{1i} Y_{t-i} + \sum_{i=1}^p \gamma_{1i} X_{t-i} + \varepsilon_{1t} \quad (2)$$

$$X_t = \alpha_2 + \sum_{i=1}^p \beta_{2i} X_{t-i} + \sum_{i=1}^p \gamma_{2i} Y_{t-i} + \varepsilon_{2t} \quad (3)$$

where ε_{1t} and ε_{2t} are white noise error terms that may be contemporaneously correlated.

2.3 Testing Procedure

Testing whether X Granger-causes Y is equivalent to testing whether the coefficients γ_{1i} in Equation (2) are jointly equal to zero:

$$H_0 : \gamma_{11} = \gamma_{12} = \dots = \gamma_{1p} = 0 \quad (4)$$

This hypothesis can be tested using an F -test (or equivalently, a Wald test). Under the null hypothesis of no Granger causality, the test statistic follows an $F(p, T - 2p - 1)$ distribution, where T is the sample size and p is the lag order.

The `grangersearch` package performs this test in both directions, providing a complete picture of the predictive relationships between the two series.

2.4 Interpretation and Limitations

The interpretation of Granger causality results requires careful consideration of several important caveats. Most fundamentally, Granger causality is a statistical concept based on predictability rather than a philosophical statement about causal mechanisms. A finding that X Granger-causes Y indicates that lagged values of X provide statistically significant predictive information about Y beyond what is contained in the history of Y itself. However, such a relationship may arise from confounding variables, reverse causality operating at different time scales, or other non-causal processes (Peters et al., 2017). Researchers should therefore interpret Granger-causal relationships as evidence of predictive precedence rather than definitive proof of causal influence.

The standard Granger causality test relies on the assumption that the time series under analysis are stationary. When this assumption is violated, test statistics may not follow their assumed distributions, leading to spurious inference. Non-stationary series should be transformed through differencing or, when appropriate, analyzed within a cointegration framework that accounts for long-run equilibrium relationships (Toda and Yamamoto, 1995).

Results from Granger causality tests can exhibit sensitivity to the choice of lag order p in the VAR specification. Selecting too few lags may fail to capture the true dynamic relationship, while selecting too many lags reduces statistical power and may introduce spurious correlations. Information criteria such as AIC and BIC provide data-driven guidance for lag selection, though examining results across multiple reasonable lag specifications is advisable to assess robustness (Lütkepohl, 2005).

Finally, standard Granger causality tests are inherently bivariate in nature. When multiple variables are potentially relevant, pairwise tests may fail to detect relationships that only manifest when conditioning on additional variables, or may identify spurious relationships that disappear in a fuller multivariate specification. Researchers analyzing complex systems should consider whether bivariate analysis adequately captures the relevant dynamics or whether multivariate approaches are warranted.

3 Package Description

3.1 Installation

The `grangersearch` package can be installed from GitHub:

```
1 # Install from GitHub
2 devtools::install_github("nkorf/grangersearch")
```

3.2 Main Function

The primary function is `granger_causality_test()`, which accepts two time series and returns a structured result object:

```

1 granger_causality_test(.data = NULL, x, y,
2                        lag = 1, alpha = 0.05, test = "F")

```

Arguments:

- `.data`: Optional data frame or tibble containing the time series
- `x, y`: Numeric vectors or column names (if `.data` provided)
- `lag`: Integer lag order for the VAR model (default: 1)
- `alpha`: Significance level for hypothesis testing (default: 0.05)
- `test`: Test type, currently only "F" supported

3.3 Tidyverse Integration

A key design goal of `grangersearch` is seamless integration with the tidyverse ecosystem (Wickham et al., 2019). The package supports:

Pipe operators: Both the magrittr pipe (`%>%`) and the native R pipe (`|>`) are supported:

```

1 # Using native pipe
2 df |> granger_causality_test(price, volume)
3
4 # Using magrittr pipe
5 df %>% granger_causality_test(price, volume)

```

Non-standard evaluation: Column names can be passed unquoted when using a data frame:

```

1 # Unquoted column names
2 granger_causality_test(df, price, volume)
3
4 # Equivalent to
5 granger_causality_test(x = df$price, y = df$volume)

```

Broom compatibility: The package provides `tidy()` and `glance()` methods following the conventions of the broom package (Robinson, 2014):

```

1 result <- granger_causality_test(x = x, y = y)
2
3 # Tidy output (one row per direction)
4 tidy(result)
5
6 # Model summary (one row)
7 glance(result)

```

3.4 Output Structure

The function returns an S3 object of class `granger_result` containing:

3.5 S3 Methods

The package provides intuitive `print()` and `summary()` methods:

```

1 result <- granger_causality_test(x = x, y = y)
2 print(result)

```

Component	Description
<code>x_causes_y</code>	Logical: does X Granger-cause Y ?
<code>y_causes_x</code>	Logical: does Y Granger-cause X ?
<code>p_value_xy</code>	p -value for $X \rightarrow Y$ test
<code>p_value_yx</code>	p -value for $Y \rightarrow X$ test
<code>test_statistic_xy</code>	F -statistic for $X \rightarrow Y$ test
<code>test_statistic_yx</code>	F -statistic for $Y \rightarrow X$ test
<code>lag</code>	Lag order used
<code>alpha</code>	Significance level used
<code>n</code>	Number of observations

Table 1: Components of the `granger_result` object

Granger Causality Test
=====

Observations: 100, Lag order: 1, Significance level: 0.050

$x \rightarrow y$: x Granger-causes y ($p = 0.0001$)
 $y \rightarrow x$: y does not Granger-cause x ($p = 0.4521$)

3.6 Exhaustive Pairwise Search

A key feature of `grangersearch` is the ability to perform exhaustive pairwise Granger causality tests across multiple time series. The `granger_search()` function tests all directed pairs of variables in a dataset:

```
1 granger_search(.data, ..., lag = 1, alpha = 0.05, test = "F",
2                 include_insignificant = FALSE)
```

Arguments:

- `.data`: Data frame or tibble containing the time series
- `...`: Optional column selection (if omitted, all numeric columns are used)
- `lag`: Integer lag order or vector of lags to test
- `alpha`: Significance level for filtering results
- `include_insignificant`: Logical; if `FALSE`, only significant relationships are returned

For n variables, the function tests all $n(n - 1)$ directed pairs and returns results sorted by p -value. This enables rapid discovery of predictive relationships in multivariate datasets. When a vector of lags is provided, the function tests each lag order and returns the best result (lowest p -value) for each pair.

The results can be visualized as a causality matrix using the `plot()` method, which produces a two-panel heatmap display similar to correlation matrices. Figure 1 shows an example output. The left panel (Row \rightarrow Column) tests whether each row variable Granger-causes each column variable, while the right panel (Column \rightarrow Row) shows the reverse direction. Cells are colored blue for significant relationships (at the specified α level) and gray for non-significant relationships, with actual p -values displayed in each cell.

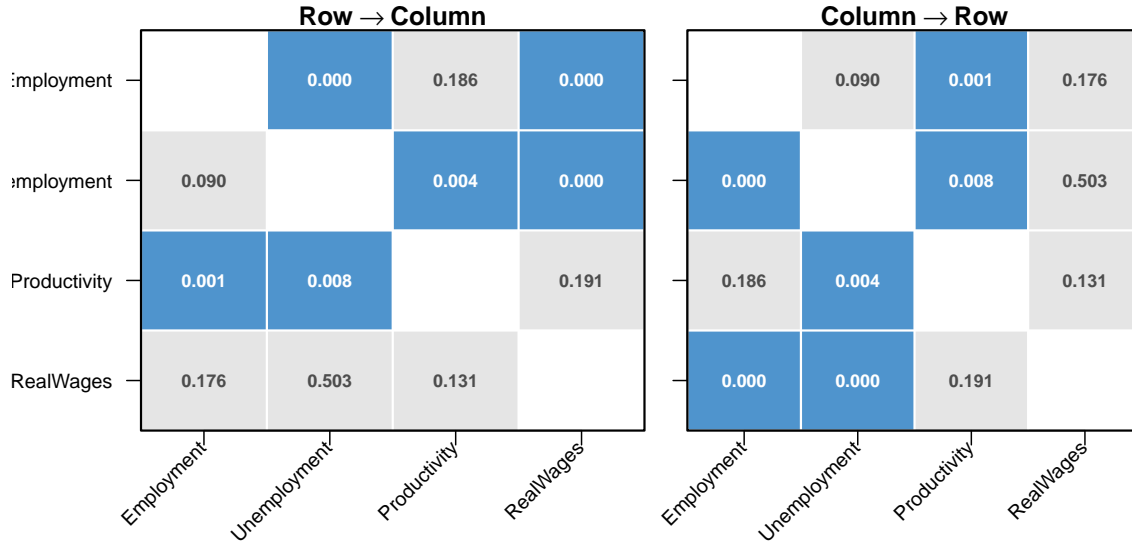


Figure 1: Causality matrix visualization from `plot(granger_search())` applied to the **Canada** dataset from the **vars** package (Pfaff, 2008). This dataset contains quarterly Canadian macroeconomic data (1980–2000) with four variables: Employment, Productivity, Real Wages, and Unemployment. The left panel (Row → Column) tests whether each row variable Granger-causes each column variable; the right panel shows the reverse. Cells are colored blue for statistically significant relationships ($p < 0.05$) and gray otherwise, with p -values displayed. Notable findings include Employment strongly Granger-causing Unemployment ($p < 0.001$), consistent with labor market dynamics.

3.7 Lag Order Selection

Choosing the appropriate lag order is critical for Granger causality testing. The `granger_lag_select()` function systematically evaluates multiple lag orders to help identify the optimal specification:

```
1 granger_lag_select(.data, ..., lag = 1:4, alpha = 0.05, test = "F")
```

The function returns an S3 object of class `granger_lag_select` containing detailed results for each lag order tested. A `plot()` method provides visualization of how p -values and test statistics vary across lag orders:

```
1 # Analyze lag selection
2 lag_analysis <- df |> granger_lag_select(x, y, lag = 1:8)
3
4 # Visualize results
5 plot(lag_analysis)
```

The visualization displays p -values across lag orders for each direction of causality, with a horizontal reference line at the significance level. This helps researchers assess the sensitivity of results to lag specification and identify robust causal relationships. Figure 2 shows an example output from `plot(granger_lag_select())`, illustrating how p -values for both causal directions vary across lag orders 1 through 8.

4 Examples

4.1 Basic Usage

The following example demonstrates the package using simulated data where X Granger-causes Y :

Granger Causality: P-values by Lag Order

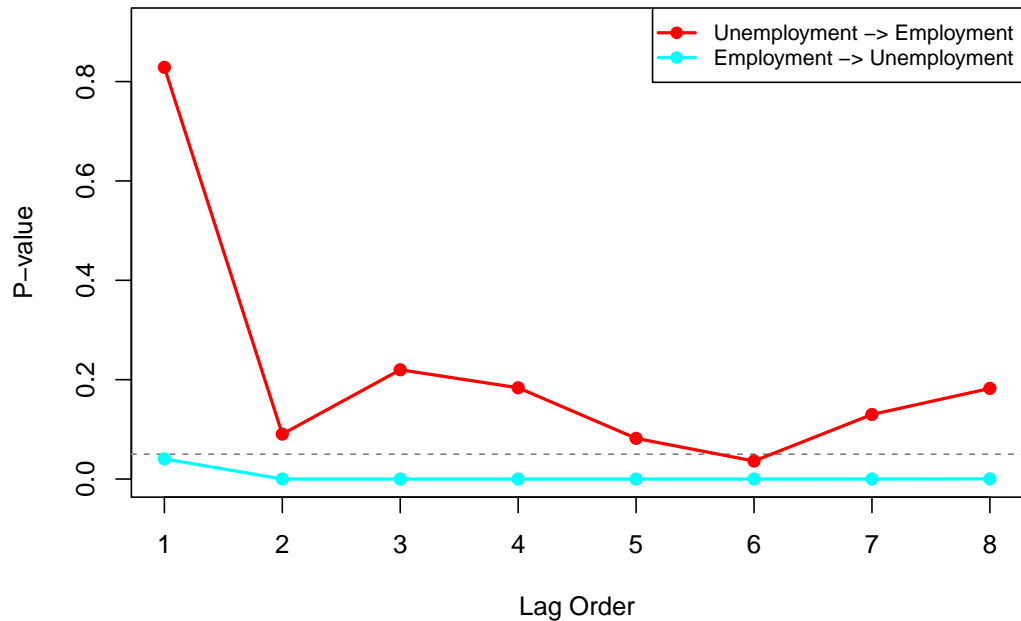


Figure 2: Lag selection analysis for Employment and Unemployment from the **Canada** dataset. The plot displays p -values for both directions of Granger causality across lag orders 1 to 8. The horizontal dashed line indicates the significance threshold ($\alpha = 0.05$). Employment consistently Granger-causes Unemployment across all lag specifications (solid line below threshold), while the reverse relationship remains insignificant at most lags, demonstrating a robust unidirectional causal relationship consistent with labor market theory.

```

1 library(grangersearch)
2
3 # Generate data with known causal structure
4 set.seed(123)
5 n <- 200
6 x <- cumsum(rnorm(n)) # Random walk
7 y <- c(0, 0.7 * x[1:(n-1)]) + rnorm(n, sd = 0.5) # Y depends on lagged
8   X
9
10 # Perform test
11 result <- granger_causality_test(x = x, y = y)
12 print(result)

```

Granger Causality Test

=====

Observations: 200, Lag order: 1, Significance level: 0.050

x -> y: x Granger-causes y (p = 0.0000)
y -> x: y does not Granger-cause x (p = 0.8934)

As expected, the test correctly identifies that X Granger-causes Y but not vice versa.

4.2 Using with Data Frames

The tidyverse-style interface enables clean, readable code:

```
1 library(tibble)
2
3 # Create data frame
4 df <- tibble(
5   price = cumsum(rnorm(200)),
6   volume = c(0, 0.5 * price[1:199]) + rnorm(200)
7 )
8
9 # Test using pipe and column names
10 result <- df |>
11   granger_causality_test(price, volume)
12
13 # Get tidy results
14 tidy(result)
```



```
# A tibble: 2 x 6
  direction      cause effect statistic p.value significant
  <chr>          <chr> <chr>      <dbl>   <dbl> <lgl>
1 price -> volume price  volume    45.2 1.23e-9 TRUE
2 volume -> price volume price      0.4 5.34e-1 FALSE
```

4.3 Comparing Different Lag Orders

The lag order can affect results. Researchers should explore multiple specifications:

```
1 # Test with different lags
2 results <- lapply(1:4, function(p) {
3   result <- granger_causality_test(x = x, y = y, lag = p)
4   data.frame(lag = p,
5             p_value_xy = result$p_value_xy,
6             p_value_yx = result$p_value_yx)
7 })
8 do.call(rbind, results)
```

4.4 Example Dataset

The package includes an example dataset `example_causality` with known causal structure:

```
1 data(example_causality)
2
3 # Inspect data
4 head(example_causality)
5
6 # Run test
7 example_causality |>
8   granger_causality_test(cause_x, effect_y) |>
9   summary()
```

4.5 Exhaustive Search Example

The `granger_search()` function enables discovery of causal relationships across multiple variables. The following example demonstrates searching for Granger-causal relationships in a dataset with four economic variables:


```

1 library(grangersearch)
2
3 # Simulate economic data with known structure
4 set.seed(42)
5 n <- 200
6 econ_data <- tibble(
7   gdp = cumsum(rnorm(n)),
8   unemployment = c(0, -0.3 * gdp[1:(n-1)]) + rnorm(n, sd = 0.5),
9   inflation = cumsum(rnorm(n)),
10  interest_rate = c(0, 0.5 * inflation[1:(n-1)]) + rnorm(n, sd = 0.3)
11 )
12
13 # Search all pairwise relationships
14 results <- econ_data |> granger_search()
15 print(results)

```

Granger Causality Search Results

=====

4 significant relationships found (alpha = 0.05)

cause	effect	p.value	lag	significant
gdp	unemployment	0.0001	1	TRUE
inflation	interest_rate	0.0002	1	TRUE
...				

The search correctly identifies the two planted causal relationships: GDP Granger-causes unemployment, and inflation Granger-causes the interest rate.

4.6 Lag Selection Analysis

The `granger_lag_select()` function helps identify the optimal lag order and assess result robustness:

```

1 # Analyze how results change with lag order
2 lag_results <- econ_data |>
3   granger_lag_select(gdp, unemployment, lag = 1:6)
4
5 # View detailed results
6 print(lag_results)
7
8 # Visualize p-values across lag orders
9 plot(lag_results)

```

The plot shows p -values for both directions of causality (GDP \rightarrow unemployment and unemployment \rightarrow GDP) across different lag orders. Consistent significance across multiple lag specifications provides stronger evidence for a robust causal relationship.

4.7 Combining Search and Lag Selection

For comprehensive exploratory analysis, the exhaustive search can be combined with multiple lag specifications:

```

1 # Search with multiple lags (returns best lag per pair)
2 results <- econ_data |>
3   granger_search(lag = 1:4)
4

```

```
5 # Results include optimal lag for each relationship
6 print(results)
```

This approach automatically selects the lag order that yields the strongest evidence for each potential causal relationship, while still controlling for multiple testing considerations.

5 Discussion

5.1 Practical Recommendations

The application of Granger causality testing requires careful attention to several methodological considerations. First and foremost, researchers should verify that time series are stationary before conducting tests. Unit root tests such as the Augmented Dickey-Fuller (ADF) test or the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test should be applied to each series, and non-stationary series should be differenced appropriately before analysis (Toda and Yamamoto, 1995).

The selection of lag order represents another critical decision that can substantially influence results. Information criteria such as AIC and BIC provide guidance for lag selection, and the `VARselect()` function in the `vars` package offers a convenient implementation. The `granger_lag_select()` function introduced in this package provides an alternative approach by allowing researchers to visualize how test results vary across different lag specifications. Results that remain significant across multiple reasonable lag orders provide stronger evidence for robust causal relationships.

Beyond statistical considerations, researchers should ground their analyses in substantive theory. Statistical significance alone does not establish meaningful relationships; the identified patterns should be interpretable within the relevant theoretical framework. Furthermore, it is essential to maintain appropriate caution when interpreting results. Granger causality fundamentally measures predictive relationships rather than causal mechanisms in the philosophical sense (Peters et al., 2017). A finding that variable X Granger-causes variable Y indicates that past values of X improve predictions of Y , but does not preclude the possibility that both variables are driven by an unobserved common cause.

5.2 Comparison with Alternatives

The R ecosystem offers several packages for Granger causality analysis, each with distinct strengths. The `vars` package (Pfaff, 2008) provides comprehensive VAR modeling capabilities including a `causality()` function for testing Granger causality. This package offers extensive functionality for advanced users who require fine-grained control over model specification and diagnostics. The `lmtest` package (Zeileis and Hothorn, 2002) provides a `grangertest()` function that offers a streamlined interface for bivariate testing. Additionally, specialized packages such as `bruinger` cater to domain-specific applications in neuroimaging.

The `grangersearch` package occupies a distinct niche within this ecosystem. Its primary contribution lies in the combination of exhaustive search capabilities with a tidyverse-compatible interface. While existing packages require users to specify variable pairs manually, `grangersearch` automates the discovery process across multiple variables, making it particularly suitable for exploratory analysis of multivariate datasets. The integration with tidyverse conventions through pipe operators and non-standard evaluation ensures that the package fits naturally into modern R workflows, while the structured output objects and broom compatibility facilitate reproducible research practices.

5.3 Limitations and Future Work

Despite its comprehensive functionality for pairwise Granger causality analysis, the **grangersearch** package has several limitations that merit acknowledgment. The current implementation conducts tests on a pairwise basis, which does not account for potential confounding by other variables in the dataset. True multivariate Granger causality testing, which conditions on all other variables simultaneously, would provide more robust inference but requires substantially more complex implementation.

The package currently implements only the F -test for hypothesis testing. While this approach is standard and widely accepted, alternative test statistics such as Wald tests and likelihood ratio tests offer different properties that may be preferable in certain applications. Additionally, the package assumes that input series are stationary and does not provide built-in handling for cointegrated systems, which require specialized treatment through error correction models.

When conducting exhaustive searches across many variable pairs, the multiple testing problem becomes salient. Users should consider adjusting significance levels using procedures such as Bonferroni correction or false discovery rate control to maintain appropriate error rates. Future versions of the package may incorporate built-in multiple testing corrections, automatic stationarity diagnostics with differencing recommendations, conditional multivariate testing capabilities, and additional test statistics to address these limitations.

6 Conclusion

The **grangersearch** package provides a user-friendly interface for Granger causality testing and discovery in R. By combining statistical rigor with modern R programming practices, the package makes this important econometric technique accessible to a broader audience of researchers and practitioners.

Key contributions include: (1) exhaustive pairwise search functionality enabling automated discovery of causal relationships across multiple variables, (2) lag order optimization tools with visualization to support robust model specification, (3) tidyverse integration ensuring the package fits naturally into modern data analysis workflows, and (4) structured output objects that facilitate programmatic use and reproducible research.

The package is intended to be useful for researchers investigating predictive causal relationships in time series data across economics, finance, neuroscience, and other disciplines where understanding temporal dependencies between variables is of interest.

Computational Details

The **grangersearch** package (version 0.1.0) is available from GitHub at <https://github.com/nkorf/grangersearch>. It requires R version 4.1.0 or higher and depends on the **vars**, **rlang**, **tibble**, and **generics** packages. All examples were run using R version 4.5.1.

Acknowledgments

The author thanks the developers of the **vars** package (Pfaff, 2008), which provides the underlying VAR modeling infrastructure used by **grangersearch**. The tidyverse ecosystem (Wickham et al., 2019) and **rlang** package (Henry and Wickham, 2024) also contributed to the design philosophy and implementation of the package.

References

- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.
- Henry, L. and Wickham, H. (2024). *rlang: Functions for Base Types and Core R and 'Tidyverse' Features*. R package version 1.1.4.
- Hiemstra, C. and Jones, J. D. (1994). Testing for linear and nonlinear Granger causality in the stock price-volume relation. *The Journal of Finance*, 49(5):1639–1664.
- Kraft, J. and Kraft, A. (1978). On the relationship between energy and GNP. *The Journal of Energy and Development*, 3(2):401–403.
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.
- Peters, J., Janzing, D., and Schölkopf, B. (2017). *Elements of causal inference: Foundations and learning algorithms*. MIT Press.
- Pfaff, B. (2008). *vars: VAR Modelling*. R package version 1.6-1.
- Robinson, D. (2014). broom: An R package for converting statistical analysis objects into tidy data frames. *arXiv preprint arXiv:1412.3565*.
- Seth, A. K., Barrett, A. B., and Barnett, L. (2015). Granger causality analysis in neuroscience and neuroimaging. *Journal of Neuroscience*, 35(8):3293–3297.
- Sims, C. A. (1972). Money, income, and causality. *The American Economic Review*, 62(4):540–552.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1):1–48.
- Toda, H. Y. and Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66(1-2):225–250.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686.
- Zeileis, A. and Hothorn, T. (2002). *lmtest: Testing Linear Regression Models*. R package version 0.9-40.