Algebraic network model identification for inflation dynamics forecasting

Fredy Vides[a], Carolina Ruiz[a], Aníbal Rojas[a], Esdras Josiel Sánchez[a]

[a] *National Commission of Banks and Insurance Companies, Tegucigalpa, 11101, Honduras*

1. Introduction

In this document, an algebraic network model identification method for inflation dynamics forecasting is presented. The modeling approach proposed in this document considers not just the local variables often related with inflation rate dynamics in a given country or economy, but also the corresponding dynamic variables observed in other countries with significant correlation with the inflation processes of the country under consideration, forming a network whose dynamics can then be approximately identified using available training inflation dynamics data. The identified network dynamics is then used as an initial predictor model, whose predictive accuracy can then be improved by identifying a corrector model based on validating data sampled from the inflation process under consideration, and allow the interpretation of the overall inflation process dynamics as a closed-loop control system. After the control system corresponding to the inflation process has been identified, both the predictability of future inflation rates and the reachability of some inflation expectations can be estimated for some suitable confidence intervals.

The first step of the inflation forecasting approach proposed in this document, is determined by the identification of a network of the countries, internal or external factors with the most significant (linear) contributions to the inflation related processes in a particular country under consideration.

As observed by Capistrán and Ramos-Francia [1], autoregressive models can be effectively used to estimate important indicators for inflation dynamics identification, and in the work reported by Silva and Piazza in [2] it can be observed that a vast amount of machine learning methods can be successfully implemented for inflation forecasting. In this study we propose a multi-variate network based technique for algebraic model identification of inflation dynamics, that builds on the method presented in [3] to identify sparse representations of algebraic models that can be used for inflation forecasting.

For the methodology proposed in this document, the node **1** always represents the country whose inflation process dynamics is the main subject under study. This distinction is important as the the network structure and configuration may change over time as the contribution of some countries, internal or external elements become less relevant for predictions corresponding to some time horizon. The relational network identification method proposed for this study will be based data-driven linear dependence measurements based on standard paired correlations estimates.

Once a relational network $\mathscr{N}(\mathbf{c})$ for the inflation process of a given country $\mathbf{c}$ under consideration, has been identified. Given a dynamic financial variable vector v_t determined by the current network configuration $\mathscr{N}(\mathbf{c})$, that can be observed/measured directly, and a dynamics indicator (variable) vector h_t related to v_t that can not be observed/measured directly, where t is an arbitrary time step, we approach the model identification problem as a parametric system identification corresponding to the estimating the following transformations.

$$\xrightarrow{e_t(v)} \boxed{\mathbf{P}_T(\mathbb{L},\mathbf{p})} \xrightarrow{v_t} \boxed{\mathbf{C}_T(\mathbf{q})} \xrightarrow{h_t} \tag{1.1}$$

For the study reported in this document, the time scale for the discrete time step t corresponds to years.

Once the left block in (1.1) has been identified, based on some training data measured from the elements represented in the relational network $\mathscr{N}(\mathbf{c})$ under consideration, a forecast variable vector $\hat{h}_t$ can be estimated using the predictive model whose output is determined by the expression:

$$\hat{h}_t = \mathbf{C}_T(\mathbf{q})v_t \tag{1.2}$$

where the variable v_t is determined by the algebraic model

$$\left(I - \sum_{j=1}^{P} \mathbf{\Phi}_j(T,\mathbf{p})\mathbb{L}^j\right) v_t = e_t(v). \tag{1.3}$$

with $\mathbb{L}$ denoting the lag operator determined the expression $\mathbb{L}^j v_t = v_{t-j}$ for each positive integer j. The entries of the matrix coefficients of the polynomial

$$\mathbf{\Phi}_T(\mathbb{L},\mathbf{p}) := \sum_{j=1}^{P} \mathbf{\Phi}_j(T,\mathbf{p})\mathbb{L}^j$$

are determined by identifying the matrix of parameters $\mathbf{p}$ using the aforementioned training data set. The matrix of parameters $\mathbf{q}$ corresponding to the block $\mathbf{C}_T(\cdot)$ are identified using a validating data set also measured from the elements represented in the network $\mathscr{N}(\mathbf{c})$.

2. Methodology

In order to identify the predictive models outlined in §1 we will consider decompose the modeling process into two main tasks: Relational network identification and switching system identification on the relational network.

2.1. Algebraic network models

2.1.1. Relational network identification

Given a country of interest $\mathbf{c}$ represented by the $\mathbf{1}$ node on a relational network $\mathscr{N}_T(\mathbf{c})$ that remain valid for a set of time steps $\tau, \tau+1, \ldots, \tau+T$ corresponding to a time period of T years after the reference time τ. The relational network $\mathscr{N}_T(\mathbf{c})$ is determined in terms of the entries of the correlation matrix $\mathbf{C}_v(T)$ corresponding to the vector signal $\{v_\tau, \ldots, v_{\tau+T}\}$, with absolute values above a certain threshold $0 < \varepsilon < 1$ chosen in a judicious manner.

2.1.2. Algebraic predictor identification

The algebraic predictor block $\mathbf{P}_T(\mathbb{L}, \mathbf{p})$ of the model proposed in this document is identified by solving the optimization problem determined by the expression

$$\mathbf{p} = \arg \min_{\hat{\mathbf{p}} \in \mathbb{R}^{n \times nP}} \sum_{t=0}^{T} \|(I - \mathbf{\Phi}_T(\mathbb{L}, \hat{\mathbf{p}}))\, v_{\tau+t}\|_F^2, \qquad (2.1)$$

for some suitable training data $\Sigma_{\mathscr{T}} := \{v_\tau, \ldots, v_{\tau+T}\}$. An approximate solution to the optimization problem (2.1) can be computed combining Ridge regression techniques with the sparse representations methods presented in [3].

2.1.3. Algebraic corrector identification

The algebraic corrector block $\mathbf{C}_T(\mathbf{q})$ of the model proposed in this document is identified by solving the optimization problem determined by the expression

$$\mathbf{q} = \arg \min_{\hat{\mathbf{q}} \in \mathbb{R}^{n \times 1}} \sum_{t=0}^{S} \left\| \tilde{h}_{\tau+t} - \hat{\mathbf{q}}^\top \mathbf{\Phi}_T(\mathbb{L}, \mathbf{p}) \tilde{v}_{\tau+t} \right\|_F^2, \qquad (2.2)$$

for some suitable validating data set $\Sigma_{\mathscr{V}} := \{(\tilde{h}_\tau, \tilde{v}_\tau), \ldots, (\tilde{h}_{\tau+T}, \tilde{v}_{\tau+T})\}$. An approximate solution to the optimization problem (2.2) can be computed combining Ridge regression techniques with constrained nonnegative least squares methods.

3. Computational Experiment: Inflation Rates Dynamics in Honduras

3.1. Algebraic network model

Applying the ideas presented in §2.1 one can obtain predictions corresponding to two potential scenarios for inflation evolution in Honduras illustrated in Figure 1, that correspond to two different control system identification hypotheses.

The global and local relational networks corresponding to the training inflation rates dynamics data under consideration, along with the matrix representation of the algebraic VAR method together with its corresponding spectrum, are illustrated in Figure 2.

The blue line in the bottom-right plot in Figure 2 represents the unit circle.

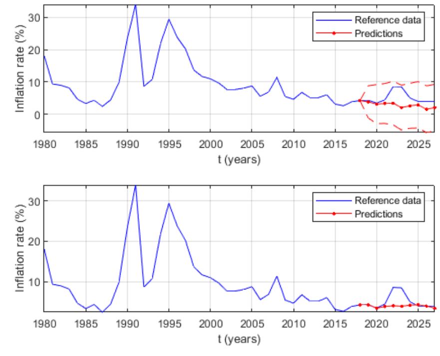

Figure 1: Predictions corresponding to two potential scenarios for inflation rates evolution in Honduras using the proposed algebraic system identification approach. AR type model (top) and VAR type model (bottom).

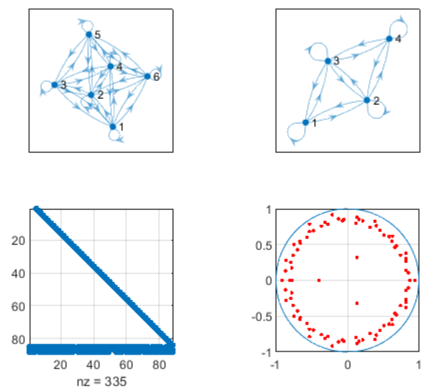

Figure 2: Global (top-left) and local (top-right) relational networks. Node assignments: 1: Honduras, 2: USA, 3: Guatemala, 4: El Salvador, 5: Central America, 6: Costa Rica. Matrix representation $\mathbf{P}$ of the algebraic predictor (bottom-left). $\sigma(\mathbf{P})$ (bottom-right).

4. Data sources

The data sources considered for this study are time series data sets corresponding to dynamic variables related to inflation processes, and recorded as part of the archival material from the IMF Archives. The Matlab programs used to obtain the results presented in §3 are included as part of the supplementary materials of this document.

References

[1] C. Capistrán, M. Ramos-Francia, Inflation dynamics in latin america, Contemporary Economic Policy 27 (3) 349–362. doi:https://doi.org/10.1111/j.1465-7287.2008.00128.x.

[2] Machine learning methods for inflation forecasting in brazil: New contenders versus classical models, Latin American Journal of Central Banking 4 (2) (2023) 100087. doi:https://doi.org/10.1016/j.latcb.2023.100087.

[3] F. Vides, Computing semilinear sparse models for approximately eventually periodic signals, IFAC-PapersOnLine Elsevier 55 (2022) 205–210.